TRUTH AND FALSEHOOD IN SCIENCE AND THE ARTS

Edited by
Barbara Bokus, Ewa Kosowska

TRUTH AND FALSEHOOD
IN SCIENCE
AND THE ARTS

TRUTH AND FALSEHOOD IN SCIENCE AND THE ARTS

Edited by
Barbara Bokus, Ewa Kosowska

Barbara Bokus ORCID 0000-0002-3048-0055
Ewa Kosowska ORCID 0000-0003-4994-1517

Contents

Ewa Kosowska
ORCID 0000-0003-4994-1517
Faculty of the Humanities, University of Silesia in Katowice

Barbara Bokus
ORCID 0000-0002-3048-0055
Faculty of Psychology, University of Warsaw
»Artes Liberales« Academy

Beyond Truth and Falsehood

Science and art, two areas that started being separated deliberately and consistently on the threshold of modernism, were brought closer together again in postmodernism – thanks to the humanities and social sciences. However, this growing closeness was not a simple reversal of the situation in which the split had occurred; the rules were different, the cause was different, and the objective was different. The perception of science and the arts as areas of different kinds of cultural practice appeared prominently in the thinking of Descartes, whose argument in *Discourse on Method* sought to empower science, to turn it into a separate area of human activity and to indicate the rules according to which scientific, i.e. cognitive, thinking can be distinguished from artistic, i.e. creative, thinking (Descartes, 1980). The negation of the Cartesian paradigm, in turn, coincided with a trend in linguistics that began strongly highlighting the fact that proposals publicizing the results of someone's deliberation on epistemological procedures can appear in the social sphere, but only if they are verbalized first. This observation, however, did not translate directly into artistic practice at the time. On the contrary, on the eve of the 20th century stronger emphasis was put on the idea that knowledge about the world which is transferred through a work of art originates from the pre-discursive, symbolic and archetypal realm referring to a universal pre-experience, and therefore might be verified intuitively, either individually or collectively (the issue of a work of art's acceptance

and popularity requires separate consideration that would include not just the psychological aspect but also, for example, the achievements of contemporary memetics). So, 20th-century disputes on the value of discursive and non-discursive learning turned their attention to language, as an integral part of human culture on one hand, and an independent medium governed by autonomous norms and rules on the other. Highlighting the fact that in language, you can only speak responsibly about linguistic reality, shifted the debate on the nature of truth and falsehood strongly toward logic, leaving unresolved issues of sensing, feeling, and the sufficiency of individual experience. Although those who spoke up for the value of the individual and subjective during this time included Georg Simmel, his proposals initially did not win the approval of academic circles and started being appreciated a little later (Simmel, 2007). The problem of the cognitive value of what is subjective and individual remained an intriguing mystery. In the late 19th century (1894) Wilhelm Windelband tried to systematize the status of procedures used for describing individual events from the past or analyzing original products (of literature, visual arts, music, intellect). He proposed a division into idiographic sciences, i.e. those reporting on the unique, and nomothetic sciences, i.e. those determining laws (Windelband, 1992). His follower Heinrich Rickert decided it would be more appropriate to distinguish between cultural (idiographic) and natural (nomothetic) sciences (Rickert, 1921). Windelband favored the idea of truth as an absolute value, independent of any pragmatic circumstances (Windelband, 2008). Rickert's stance was similar, as he gave truth the status of a value – the main object of scientific cognition. Simmel's proposals, especially his concept of "subjective culture," went beyond those distinctions and were not verifiable. Neither were they possible to falsify, i.e. undergo procedures popularized by Karl Popper (2019) for the purpose of checking "bold hypotheses" without which science could not develop. Today we can already see limitations in the paradigm of Popper's falsificationism, especially when we try to apply it to the humanities and social sciences. For example, in academic psychology there are two methodological "camps" – **quantitative**, relying as much as possible on experiments, statistical analyses and conclusions about hypotheses, and **qualitative**, aiming to learn about the subjective perspective of the individual being studied.

A subjective "sense of truth," although based on acknowledging truth's objective value, hands the power of judgment to the acting subject. However, some doubt still remains as to the possibility of properly exercising that power in a situation when the judgments themselves are expressed in language. Of course the material of language allows true and false sentences – in the logical sense – to be uttered, but it does not fulfill the requirement of absolutely true statements about the non-linguistic reality. This dilemma was already noticed earlier, and led to the validation of various approximation procedures, among which granting special status to metaphors came to the forefront. At the end of the 18th century this was a new idea, but with time, metaphorical language made its way into science and gradually became an independent medium, with a life of its own, generating its own problems. Among the most important consequences of uncovering the connection between science and the choice of linguistic procedures for the transfer of research results in both the humanities and the natural sciences, we should mention the "narratological turn" that allowed practically any scientific statement to be considered in terms of a linguistic statement, in terms of a kind of "story" about a specific case or set of cases; this could be a monophonic story or one being part of a polyphony. Thus, we can approach the scientific statements of a given time in the same way Mikhail Bakhtin (1970) approached the novels of Dostoyevsky. These novels do not represent a single, overriding vision of the author, but are the result of dialogues between independent individuals and their points of view or ideologies. Therefore, we can say after Bakhtin that truth is not born in one person's mind but comes into being between people seeking it together; it is born in a process of dialogue that inevitably is entangled in linguistic conventions.

Thus, today we can recognize that fields situated on the boundary between science and the arts include, for example, psychology's Dialogical Self Theory, which posits that human personality is composed of a number of internal voices representing different social roles, aspects or traits of a person. These voices are constantly in dialogue with one another, trying to reduce tensions and resolve contradictions, make decisions and lend meanings. Underlining the subjective nature of experience, Dialogical Self Theory says that human personality is unavoidably shaped by cultural and civilizational conditions, which today

are also shifting away from objective, timeless facts and values toward more individualized "truths."[1]

The question of truth and falsehood in science and art has thus assumed a new dimension in our times: For some in our community, the classic understanding of truth, positivist in spirit, is still the main reference point in judging phenomena and processes; for others, a guarantee of scientific character, also in scientific analyses of art, lies in concentrating our research attention on the interpretation of narrative procedures. This means one relevant question is that of **what *truth* is in relation to the human individual**. Invoking Dilthey's distinction (Dilthey, 1982, 2005), we would need to resolve whether sufficiently *explaining* human anatomy, physiology and neurobiology will bring the same effect as *understanding* a person's individual experience.

The collection of analyses and interpretations presented in this volume seems to testify to an ongoing intensive search for an answer on how the problem of truth and falsehood in science and the arts should be investigated today. The individual proposals are mostly "case studies," pointing to the opposition from the title or one of its elements in the context of specific phenomena and events. However, some of the authors offer more synthetic summaries of progressively growing knowledge on the significance of nuances in distinguishing between what is true, probable, real, natural, authentic, reliable or clear, and what is untrue, false, fake, pretend, inauthentic, stylized, ironic, illusory, deceitful, erroneous, mistaken, incorrect, unreliable or wrong. The richness of synonyms in language does not so much make it easier to support relativism in relation to Truth as it indicates that social experience and the truth of language demands that we exercise great caution when expressing absolute value judgments. Today those positivist hopes for the discovery of objective facts have

---

[1] The concept of **possible worlds** is quite a good illustration of individualized truths. Umberto Eco in *Lector in fabula* (1994, p. 190) writes that the characters within a narrated world assume propositional attitudes. For example, Little Red Riding Hood thinks the character in the bed is her grandmother (whereas, for the reader's benefit, the plot has negated the girl's belief). What the girl believes is her own b e l i e f structure, but it also remains one of the states of the *fabula*. In this way, the *fabula* proposes two states of affairs: one in which it is the wolf lying in the bed, and another in which the bed's occupant is the grandmother. We know straight away (but the girl is not aware of it until the end of the story) that one of these states is presented as being true and the other as untrue.

largely died down. Instead, the relativity of knowledge is being underlined more and more often. The replies we obtain depend to a large extent on the questions asked, and the act of observation inevitably turns the observer into an element of the whole system.

Presenting the specific proposals of our authors, we need to warn readers that this is a special collection: It is the outcome of a conference organized for the conclusion of the 10-year Inter-University Program of Interdisciplinary PhD Studies at the »Artes Liberales« Academy. It was attended by the academic teachers, who have supervised some excellent doctoral dissertations, and the program's students themselves, most of whom now hold doctoral and postdoctoral degrees. The program was headed by Professor Barbara Bokus, and it was also she who proposed a few years ago that "Master Seminars" be added to the program. The seminar for the program's jubilee, and for its closing in the present format, was chaired by Professor Jerzy Axer and preceded the conference, an outline of whose subject matter will follow.

Jerzy Axer gave his talk in the form of a biographical story. This enabled him to give an accessible presentation of complicated problems connected with the development of scientific rules of editing Classical texts. The liminal experience of confronting the rigid foundations of positivist methodology with new reception theories and narratological concepts led the Professor to pose the fundamental question of the scientific credibility and veracity of the products of interpretation as it starts moving dangerously close to the border of falsity. Axer's erudite argumentation is rounded off with the following conclusion: "*Ars critica* consists of scientific activity, artistic work and an exercise in forgery of old masters, all rolled into one. Perhaps that is what makes it so fascinating and allows it to bring so many triumphs and disappointments, depending on the temperament, character and mood of whoever is practicing it."

Contemporary *ars critica* is sometimes helpless, however, when faced with the need to authenticate a miraculously discovered document from the past. This is the case recounted by **Prze-mysław Piwowarczyk**, whose attention was drawn to a media storm, almost self-inflicted by the Harvard Divinity School, on the threshold of the 21st century. Karen King, a member of the staff, was offered a manuscript that she called the *Gospel of Jesus' Wife*. The author recounts the problems with confirming the find's

authenticity almost as a crime story. However, the scientific value of his thoughts on the veracity of newly discovered artifacts, and on the capabilities of modern-day scientific methods of verifying unexpected finds, is of unquestionable importance.

The focus of **Karol Wilczyński's** interest is the output of Al-Ġazālī, an Arab thinker from the turn of the 12th and 13th centuries, an opponent of falsafa, a philosophy that first became popular more than a century before he was born and lasted almost four hundred years. Al-Ġazālī saw falsafa as a dangerous movement, potentially breaking with Islam, abandoning religious practices, trusting in reason and, as a result, erring to an extent making it impossible to find truth and salvation. "Pride connected with wrongful practice and knowledge unaware of its own limitations ultimately leads to unbelief," and "the philosophers' undoing is their pride and their belief that science alone can be enough for their salvation and happiness." Excessive trust in the power of the mind leads people astray and away from all that is important and true. Presenting his own interpretation of selected views of Al-Ġazālī, the author draws our attention to the renaissance of interest in his ideas currently observed in Arab countries.

On the example of the theory of poetic art of Maciej Sarbiewski (1595–1640), a Polish Jesuit, **Maria Łukaszewicz-Chantry** considers the gap between assumed proof of truth when a fact is documented by a record, and poetry which expresses truth indirectly, employs metaphors and allegories, often constructs nonexistent beings, to use them as examples pointing to the existence of phenomena and regularities that go unnoticed in the process of direct perception.

Izabella Zatorska traces the relations between illusion and truth in theater from the Baroque to the Romantic period, setting out to show how Baroque and Rococo theater influenced the behavior of the elites and to what extent those elites were aware of the social consequences of intentional theatralization of life. Those consequences were definitely noticed by some writers, who proposed a secondary theatralization of the scene: "turning the game of illusion into the subject; this is a method allowing the truth to be sought – over and above the immediacy of the masquerade." This kind of transformation, effected in the time of neoclassicism and then Romanticism, did not eradicate the theatralization of life as such, but fundamentally changed its forms and scope. What remained, however, was the problem – important

in both life and theater – of lying as intentional misleading and speaking untruths, and illusion creating appearances thanks to which truth could be revealed.

Wojciech Sajkowski draws our attention to 18th- and 19th-century French-language descriptions of the Morlachs. The authors of the accounts he has chosen quite unanimously judged that "the Southern Slavs who inhabit provincial Dalmatia are simple, naturally good and sincere." The author links the convention of these narratives with a peculiar valorization of truth. This emerged in the 17th century as a result of contacts between educated, or at least literate, residents of Western Europe with "uncivilized" peoples. Among those considered uncivilized were Slavic shepherds, whose "confidence, devotion and sincerity are linked to an inability to understand abstractions and sophisticated truths of the civilized world." The opposition of civilized versus uncivilized, disseminated by Georges-Louis Leclerc de Buffon, found confirmation in the observations of travelers of the time. The writers who penned the accounts under consideration, theoretically respecting the idea of natural law, noted that in the world of civilization "yes" does not always mean "yes" and "no" does not always mean "no." In contact with that world, "people of nature" acting artlessly fall victim to their own naivety. Thus, the Slavs' honesty becomes a sign of their uncivilized character.

Adam Grzeliński points out that "classical philosophical aesthetics … developed several solutions to the problem of justifying the validity of aesthetic judgments." The main authors of such solutions in British aesthetics were Anthony Ashley Cooper Shaftesbury, Edmund Burke, and David Hume. Their achievements were synthesized in Immanuel Kant's concept of reducing the phenomenon of beauty to its simplest form and according aesthetic judgements the value of universality and validity. Grzeliński focuses on presenting the views of George Santayana (1863–1952), who offered a critique of 18th-century philosophers' efforts to establish conditions for the validity of judgements of taste; in this, he sets out to show that these seemingly polemical stances can be reconciled to some extent.

Anna Żymełka-Pietrzak highlights the significance of ideas put forward by Johann Georg Hamann, who was one of the first to negate systematic philosophy and saw the very notion of a system as "a hindrance to truth." That is why this thinker from Königsberg gave up the precise language typical of rational

discourse in favor of metaphorical disquisition introducing people to the essence of things in a way making it more easily comprehensible. The actual process of getting to the truth hidden amidst a tangle of symbols and metaphors was to be illustrated by Lichtwer's story of the many dresses of the woman-fable attacked by robbers. Stripped of her clothing, the naked fable, like the naked truth, is something "no one can bear to look at." But in Hamann's interpretation, the secret of truth is that it is bodiless, and thus removing successive layers of clothing can be a fascinating road to learning, but arriving at the goal thwarts the purpose, for there is only one Truth, and it is connected to divine reality inaccessible to humans. This nature of truth means we can only speak of it indirectly, through metaphorical language that has a necessary presence in philosophy.

In her paper "Worthless yet priceless: The truths and economics of poetry" **Marta Baron-Milian** discusses the truth of a poetic text not so much in opposition to falsehood, but rather in the context of knowledge to be extracted indirectly. The semantic potential of a poetic message and its unobvious character as a source of new meanings was not an evident value on the threshold of modernity. Contrary to Johann Wolfgang Goethe, who demanded the right to view the world through imagination (and not just scientific knowledge), Jeremy Bentham, for example, thought that "all poetry is misrepresentation," that it is not true to reality and is not useful. Tracing the discourse surrounding the worth of what is not useful and the truth of what is impossible to verify using scientific tools, the author succeeds in convincing her readers that there are multiple dimensions involved in the useful, and there is enormous cognitive value in the ambiguous.

The concept of reflection has played an important role in the quest for truth, whether in science or in the arts. **Paweł Tomczok** is interested in cognition and experience, which respectively belong to the realms of learning and aesthetics, in relation to the metaphor of mirroring, and also in "different forms of mirroring – above all in art and philosophy." He points to the fact that the metaphor of mirroring "quickly got ensnared in linguistic representation – it was words and sentences, notions and judgments that were meant to reflect reality," and then mirrors themselves became "historical mediators that take part in a complicated network made up of people and things." Therefore, they can be seen as nodes in a network "combining theory and practice,

science and aesthetics, technology and speculation." Highlighting the paradigmatic potential of mirroring, the author does not lose sight of the metaphor of words reflecting reality and the problem of truth this leads to.

In his paper on truth and probability in studies on religious thought, **Michał Rogalski** considers the logical consequences of a historian of philosophy assuming the position of a disengaged, objective, unbiased researcher. Such a tradition, originating from 19th-century German thought, reveals its inadequacy as soon as controversial issues come into play, when describing them requires showing "their mutual relationships and the argumentation networks forming as a result of their coexistence." It is hard to stay impartial when you have to judge the cohesion of someone else's argumentation involved in a dispute and being modified as the dispute progresses. Rogalski considers this theoretical problem on the example of the debate on Catholic modernism that took place in the last decade of the 19th century and in the early 20th century. "However, when this kind of description is used, questions inevitably need to be asked about the truth or probability of hypotheses, arguments, or entire argumentation strategies. Neither the category of truth nor the category of probability can be eliminated when using value judgments." Therefore, the author wonders how we are "to work with" these categories.

Juxtaposing "Hannah Arendt's Marranic evasions and the truth of her cryptotheology," **Rafał Zawisza** successfully points us to such an interpretation of the Gospel by Arendt, as an expert on the writings of Augustine of Hippo, that would allow for the simultaneous use of the hermeneutic tradition of Jews and Christians for explaining at least some of the mysteries of human existence. Thus, he accents the importance of natality as a counterbalance to mortality, and also the beginning of any human action, as well as extracting the idea that "labor, work, and action are 'intimately connected' with birth and death." Natality, Zawisza tells us, lies at the foundation of "the cryptotheology of singularity." This cryptotheology enabled Arendt to offer her distinctive reinterpretation of the message on the birth of Jesus. By highlighting Arendt's peculiar flirtation with two theological traditions in her oeuvre, the paper's author reveals a new dimension of truth hidden beneath that which is seemingly "sloppy" and imprecise.

Piotr Kałowski considers the problem from the volume's title in the context of the achievements of contemporary psychology.

First confirming the importance of psychotherapy, the author also points out that the reasons it is effective are not fully clear. Hence, he investigates how the notions of truth and falsehood apply to a situation when we know how a sequence of causes and effects proceeds, but we do not know why. With the help of methodological tools developed by narrative medicine, he indicates the importance of changes that occur during therapy in an area defined by a patient's narrative that is the result of reconciling many internal dialogues. The previously undisputed credibility of such a narrative can sometimes be traumatic, and its reorganization under the guidance of the right therapist, i.e. one who guarantees a safe relationship, can bring positive effects. This method, inspired by postmodernist ideas, could – according to the author – effectively support the much more popular cognitive behavioral therapy which, however, was developed in different historical circumstances.

Anna Milanowicz offers some thoughts on the relationship between truth and untruth in irony. From a psycholinguistic viewpoint, the author characterizes the intention behind ironic messages, seeing irony as the result of a certain type of social experience. The essay invokes the Socratic tradition related to cognition and the source of "storytelling" (building stories and relationships with the audience) in dramatic irony. The theory of Negation of Expectations in Irony (NEI) is proposed. The relationship between "ironic" untruth and lies is outlined. Finally, the ambiguity of "ironic" meanings is described from the perspective of gender binary stereotypes.

Adrianna Smurzyńska discusses a problem familiar to cognitivists, namely mentalization, or attributing mental states to others. The source of this competence has yet to be fully explained. The author focuses on simulation theory, which assumes that you can get to know others by simulating their mental states. This process can take place automatically and unthinkingly (intuitively), or "on a reflexive level" when one tries to put oneself in another's shoes. In this theory, "our own mind is treated as a tool which enables us to take another person's perspective." Identifying the mental states of others is only possible, however, when the person doing this does not have any mental disorders, since the "abilities to differentiate perspectives and separate one's own and others' mental states seem to be necessary conditions for adequate mentalization. The capability to simulate others' mental

states – to put oneself in someone else's shoes – is important but at the same time not sufficient."

Joanna Barska draws our attention to music as illustration, tracing this phenomenon from the Renaissance when music was ideologically subordinated to text and the principle of decorum. With time, composers developed a set of rhetorical-musical figures that were meant to signal specific content. On the other hand, "the 17th- and early 18th-century belief that the main purpose of music is to arouse passions and affects yielded the theory of affects that had an especially strong presence in the German-speaking world." This trend in music developing from the Renaissance aimed to evoke natural sounds and mental states as faithfully as possible with the use of means at the disposal of musicians in a given period.

Truth in music is also a matter of interest to **Karolina Kolinek-Siechowicz,** who focuses on historically informed performance. She convincingly argues not only that a present-day musician cannot credibly recreate the genuine sound of historical instruments or the way a given piece was performed, but also that neither today's performers nor today's audiences can break free of the constraints of their own culture and its aesthetic expectations.

Another aspect of truth in music is presented in the paper by **Anna Chęćka**. The author suggests two ways of pursuing such truth: in one, we consider the truth of a music score, while the other seeks "truth of performance," i.e. authenticity within performance practice. Analyzing the phenomenon of pianist Yuja Wang and her fictional persona Mei Jin from the novel *Chinese Piano*, the author points to some of the problems arising when we judge the credibility of a performance that goes beyond the standards accepted in a given period.

Meanwhile, **Piotr Słodkowski** is interested in the relationship between the truth of a painting when considered in intertextual relations with other pictorial representations, and the "truth of the matter" of a specific painting. From theoretical reflection on this distinction, the author extracts two extreme attitudes: "To Heidegger, the truth revealed in painting is the truth of things (*aletheia*), to Schapiro – a faithful representation referring to an object (*mimesis*)." But it was Derrida who believed that the idea that paintings "only refer us to other paintings was decisive for the deconstruction of the metaphysics of presence." The paper's

author, apparently more fascinated by Bjørnar Olsen's question of how things remember, looks at the work of Polish-Jewish modernist Henryk Streng and American pop-art precursor Robert Rauschenberg in order to show that "the truth of a painting as the truth of matter has the irreducible value of a historical source which – specifically – testifies to very different albeit important existential experiences."

Maciej Junkiert looks into the problem of confronting the Romantic model of literary history with the demands of contemporary times. He offers the hypothesis that the Polish model of the history of literature, which sometimes more and sometimes less consciously invokes fundamental texts of German philosophy from the turn of the 18th and 19th centuries, links history education to the idea of building/preserving national identity. Hence, national history of literature is a kind of site of memory protecting the self-knowledge of a group in danger of losing its social and intellectual ties to the past. The truth of the past affecting the truth of the present and the development of the future lay at the foundation of the paradigm in which the history of literature was shaped two hundred years ago; although that paradigm is not falsified today, its value is no longer appreciated. Developing a new paradigm is a challenge for present-day philology.

Finally, **Jan Kutnik** draws attention to the problem of the truth of museum exhibitions commemorating real places of torture and killing from World War II. A confrontation of visitors' expectations with the effects of the work of exhibition designers (studies conducted in 2016–2019) reveals different ways of understanding the veracity of testimonies of the war and the Holocaust. The author shows how changing history-based politics affected the exhibition concepts pursued by the Museum at Majdanek, points to a need to stick to historical truth, and underlines the educational value of similar museums.

The great majority of the authors discussing truth and falsehood in science and art in the present volume assign truth the status of an irreducible point of reference both in striving for elementary knowledge about the world and in seeking methods and artistic means of achieving this goal. The multilevel and multiple-aspect research presented here, conducted on material from different periods and different cultures, shows very clearly that truth and falsehood lie at the foundation of all human

motivation, choices, decisions, and behaviors. At the same time, however, it reveals that every bid to extrapolate the results of detailed studies into generalizations aimed at universalization – by the very fact of their discursivation – either subjects the discussion to the rules of formal logic or situates it outside the realm of truth and falsehood.

* * *

This e-book is the English version of the monograph published in Polish as *Prawda i fałsz w nauce i sztuce*, edited by Ewa Kosowska and Barbara Bokus. Both volumes have been published simultaneously by Wydawnictwa Uniwersytetu Warszawskiego. We are grateful to the monograph's reviewers – Professor Stefan Bednarek and Professor Stanisław Rabiej – for their kind remarks and suggestions. A sincere thank you to Joanna Dutkiewicz for her diligent proofreading of this volume.

Warsaw, Katowice
September 2019   The Editors

References

Bachtin, M. (1970). *Problemy poetyki Dostojewskiego* (Polish edition of Bakhtin, M. *Problemy poetiki Dostoevskogo*, transl. N. Modzelewska). Warszawa: Państwowy Instytut Wydawniczy.

Descartes, R. (1980). *Rozprawa o metodzie właściwego kierowania rozumem i poszukiwania prawdy w naukach* (Polish edition of *Discours de la Méthode pour bien conduire sa raison et chercher la vérité dans les sciences*, transl. T. Żeleński [Boy]). Warszawa: Państwowy Instytut Wydawniczy.

Dilthey, W. (1982). *Pisma estetyczne* (Polish edition of *Dichterische Einbildungskraft und Wahnsinn*, transl. K. Krzemieniowa). Warszawa: Państwowe Wydawnictwo Naukowe.

Dilthey, W (2005). Budowa świata historycznego w naukach humanistycznych (Polish edition of *Der Aufbau der geschichtlichen Welt in den Geisteswissenschaften*, transl. E. Paczkowska-Łagowska). Gdańsk: słowo/obraz terytoria.

Eco, U. (1994). *Lector in fabula* (Polish transl. P. Salwa). Warszawa: Państwowy Instytut Wydawniczy.

Popper, K. (2019). *Wiedza obiektywna: Ewolucyjna teoria epistemologiczna* (Polish edition of *Objective Knowledge: An Evolutionary Approach*, transl. A. Chmielewski). Warszawa: Wydawnictwo Naukowe PWN.

Rickert, H. (1921). *System der Philosophie. Erster Teil: Allgemeine Grundlegung der Philosophie*. Tübingen: Mohr Siebeck.

Simmel, G. (2007). *Filozofia kultury. Wybór esejów* (Polish edition of *Philosophische Kultur: Gesammelte Essais*, transl. W. Kunicki). Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Windelband, W. (1992). Historia a nauki przyrodnicze (Polish edition of Geschichte und Naturwissenschaft, transl. J. Miziński). In: Z. Cackowski, M. Hetmański (Eds.), *Poznanie: antologia tekstów filozoficznych* [Cognition: An Anthology of Philosophical Texts] (pp. 172–180). Wrocław: Ossolineum.

Windelband, W. (2008). Wola prawdy (przemówienie z okazji święta Uniwersytetu w Heidelbergu wygłoszone 22 listopada 1909 roku) (Polish edition of *Will to Truth*, a speech delivered on Nov. 22, 1909 for a celebration at the University of Heidelberg, transl. T. Kubalica). *Folia Philosophica* 26, 45–58.

Jerzy Axer
ORCID 0000-0001-9588-7940
Faculty of "Artes Liberales", University of Warsaw
»Artes Liberales« Academy

Between Science, Art, and Forgery: Latin Textual Criticism as a Case Example

Abstract

The methodology and practice of Latin text editing is presented in the paper as an example of a humanities scholar's work that combines scientific effort with artistic activity, including forgery of works by old masters. Part one outlines the methodological situation in textual criticism in the 1960s and 1970s, when the author was a university student and novice researcher. Next, the author considers what caused changes in editing methodology in subsequent decades and, finally, he comments on present-day debates surrounding the role of the academic editor, with a special focus on the interdisciplinary aspects of this issue.

Keywords: Latin textual criticism, forgery, academic editing, intertext, science, art

1. Methodologies of my youth

My true initiation into Latin textual criticism was Prof. Kazimierz Kumaniecki's seminar at the University of Warsaw's Institute of Classical Philology, in which I took part as a student and then as an assistant lecturer in 1967–1972. The seminar was of the highest world standard, following the methodological principles and rules set down in the classic textbook by Paul Maas (first edition 1927[1]). Kumaniecki learned his research strategy directly from Maas during his seminars in Berlin even before it became

---

[1] Paul Maas first made his treatise *Textkritik* public in 1927 (Maas, 1927); a separate edition was published by B.G. Teubner in Leipzig in 1949, and another one in 1957 with an important addition by the author, "Retrospect". An English translation appeared in 1958 (Maas, 1958).

the canon of *ars critica* for the reconstruction and publication of ancient texts and continued as such for several decades. Maas' strategy derived from the "Lachmann method"[2] from the first half of the 19th century. Its advantage over earlier attempts to develop the method's theoretical basis lies primarily in applying an extreme version of principles taken from reasoning typical of science. Obviously in connection with the success of Darwin's theory, Maas proposed a rigorous approach to building the stemma, i.e. the genealogical tree, of an ancient text's tradition. Framed within rules as economical as a science textbook, Maas' lecture became a global success after its author, fleeing Nazi persecution, settled in Oxford. To editing in the English-speaking world, this version of the Lachmann tradition brought revolutionary results.

Elsewhere, relatively soon after Kumaniecki's passing, I tried to describe the essential experience of participating in his late seminars on textual criticism[3]. We worked on Cicero's *De Oratore*, which Kumaniecki published as a Teubner edition in 1969. There is no doubt that the methodological foundation of his teaching derived from Maas' stemmatic tradition, at the same time challenging it; this was never formulated in a theory, but was constantly practiced in his editing. Very early on, Kumaniecki took into account the possibility of a horizontal tradition that would weaken the authority of stemmatic patterns. This appreciation of how widespread the phenomenon of *contaminatio* could be was only starting to emerge in textbooks (Reynolds, Wilson, 1968, latest edition 2013; West, 1973).

---

[2] Karl Konrad Friedrich Wilhelm Lachmann (1793–1851). Lachmann, a theologian by education, gradually developed a keen interest in classical philology. A follower of Christian Gottlob Heyne and other scholars seeking new methods of comparative reading of the classics, he analyzed early German literature (*Song of the Nibelungs*) as well as Latin texts. His greatest achievement was the development of a new method of textual criticism called the "Lachmannsche Methode." He strove for scientific reconstruction of a text, aimed at producing a version as close to the original as possible. He considered the most important stage in arranging the source material to be drawing a stemma (genealogical tree) based on an analysis of a text's different witnesses to determine the inner relationships among the different variants. He was greatly successful in his work on *Song of the Nibelungs*. He opened a new era in research on Propertius, publishing the first truly critical edition of this poet in 1816.

[3] Cf. J. Axer (1989), presented on Dec. 15, 1987 at an academic session organized at the University of Warsaw's Institute of Classical Philology for the 10th anniversary of the deaths of professors Kazimierz Kumaniecki and Adam Krokiewicz.

Kumaniecki never wrote a textbook himself, and his interest in the theory of *ars critica* was slight. On the other hand, he spent many decades conducting an internal dialogue with great Roman prose writers and poets, in constant competition with eminent philologist editors from the older and his own generation. He was not tempted to revise or question Maas' main theoretical assumptions. I think this was because he believed that textual criticism is, above all, a practical art in which the researcher's personality and talent are crucial, while rules should only help at the early stage of practicing this art.

Once the stemma has been sketched, it is time for *examinatio* and *emendatio*. And these appear to the editor as an infinite set of individual cases, each in need of separate consideration after earlier and correct performance of operations organizing the material. My Master was interested mainly in those of the endless problems requiring solving that opened the way to speculation requiring in-depth understanding of the historical and cultural context of a given text and, above all, grasping a given author's language and artistic technique. This means his academic passion and creative temperament directed him toward actions that would reconstruct the original text in a way going beyond what could be unequivocally justified by the juxtaposition of preserved sources.

He had his greatest success with the textual criticism of Cicero's texts, thanks to phenomenal knowledge of his language and a deep understanding of the mentality of the so-called last generation of the Roman Republic. He moved beyond the boundary of scientifically verifiable knowledge; this gave students the impression that he could travel in time and intuitively establish contact with people of the Roman elite from the time of the Republic's disintegration. This is not the place to wonder what historical and personal experiences had produced this special gift. Suffice it to say that it was largely impossible to pass on in the form of rules, recipes, and instructions. This made the textual criticism we practiced during those seminars rather like artistic training.

It did not take me long to achieve sufficient proficiency in *ars critica* to base my postdoctoral degree (habilitation) on an edition of Cicero's oration *Pro Roscio Comoedo* (Axer, 1976) and to write studies on textual and structural criticism of this work (Axer, 1976, 1979). However, I was helped by a very happy

coincidence. Poggio Bracciolini's autograph, which had been lost since the 15th century, had just been found in the deepest recesses of the Vatican Library, and its finder Prof. Augusto Campana informed Kumaniecki of this fact. He in turn managed to get a photocopy of the autograph. Prof. Lidia Winniczuk had already generously transferred her contract with Bibliotheca Teubneriana for the publication of *Pro Roscio Comoedo* to me.

Therefore, I had a considerable advantage over the editors from the previous 400 years who had used many different copies originating from one source. I had chanced upon a work for which both *recensio* and *examinatio* were extremely simple and boiled down to careful reading of one manuscript; and, of course, there was no need to draw a stemma. Success depended on the ability to reconstruct the fragmentarily preserved text of Cicero's defense speech as reasoning and as a legal proceeding. This script then had to be superimposed on the very distinctive community of Roman theater actors and producers; the suspicion also existed that Cicero had used the language of that community to play a literary game, seeing as he was defending an actor and the accused was the owner of a theater school. Once I had arrived at opinions on all these matters, it was time to decide about the extent to which the preserved text required correcting (which in the language of textual criticism is known as applying conjecture and emendation) – so as to achieve the highest possible probability that the text proposed by the editor had come from the pen, or – rather – stylus, of Marcus Tullius Cicero. It was the dream opportunity for intuitive and artistic textual criticism[4]. The success went to my head, I signed seven contracts with the world's leading publishers for further editions of a number of Cicero's different orations; I had every year from 1977 to 1995 all planned.

I never finished any of those editions, even though I spent hundreds of hours working on the texts. Why? Because I lost faith.

---

[4] Numerous, sometimes excessively positive reviews: Martin ven den Bruwaene, *L'Antiquite Classique*, vol. XLVI, 1977, p. 640; Heikki Solin, *Arctos* XIII, 1979; Guglielmo Ballaira, *Giornale Italiano di Filologia* VII, 1977; Raphael Palmerini, *Latinitas*, 1977; Michael Winterbottom, *The Classical Review*, NS, Vol. XXVII No. 1, 1978; Elżbieta Olechowska, *Revue des Etudes Latines*, vol. LV; Fabio Cupaiuolo, *Bollettino di Studi Latini*, Anno VII-Fasc. III, 1978; Carl Joachim Classen, *Museum Africum*, Nigeria, 1978.

2. How did I lose faith?

Maas' textbook attempted to rework Lachmann's revolutionary discoveries from the early 19th century in such a way as to guarantee that textual criticism would, in terms of scientificity, stand up to the natural sciences[5]. Working in times of rapid development of knowledge about the Earth's past, especially paleontology and geology, Lachmann – a contemporary of Georges Cuvier and Jean-Baptiste de Lamarck – had designed procedures enabling the genealogical trees of early texts to be built (he started from the Bible, later moving on to Old German and Latin). This theory continued to be improved; at the same time, Darwinism turned out to be a very strong stimulus encouraging the propagation of stemmatic diagrams distinguishing manuscript families and subfamilies, just like animal genera and species are distinguished. At the height of this way of thinking's triumph, Paul Maas decided that true science requires short and simple rules; rules that are absolutely incontrovertible and universally binding. It was his ambition to strip textual criticism of subjectivity, the editor's influence on the object of research, and any speculations that represented risky hypotheses. Kumaniecki's late seminars, described earlier, provided proof that new narratological trends, theory of reception and, finally, the gradual decline of positivist authority, undermined the thus understood scientificity of textual criticism theory, even if, technically, Maas was the methodological patron of our Master's work in this field[6].

My loss of faith, therefore, was a natural realization of this, a gradually growing awareness that the status of everything I found fascinating and appealing about textual criticism was very ambiguous from a scientific point of view. Maas was no longer excuse enough. I came to suspect I was practicing an art, not a science, and perhaps even resorting to forgery.

What did I do? First and foremost, I tried to write a textbook that would serve as an "updated Maas." I wanted to

---

[5] Of course the beginnings of Lachmann's stemmatic strategy can already be found in the work of great 15th- and 16th-century philologists. However, fundamental changes in the way textual criticism was practiced only took place thanks to the work and authority of Lachmann himself.

[6] Apart from myself, this seminar was also attended by Elżbieta Olechowska, an excellent editor of Cicero's orations *Pro CN. Plancio* and *Pro Rabirio Postumo*, B.G. Teubner, Stuttgart 1981.

restore harmony between scientific procedure and freedom to interpret and offer textual interpretation ideas with a resultant impact on a text's reconstruction; between scientific procedure in the analysis of sources and freedom to hypothesize. I did not publish any "new Maas," but tested my position by communicating it to classical philologists and researchers of antiquity representing different schools of textual criticism. I presented it in Germany (including Heidelberg, Tübingen, Munich), Italy (Rome, Naples, Turin), Spain (Madrid, Cadiz, Barcelona), the United States (Northwestern University, Harvard) as well as Centro Ciceroniano congresses. This manifesto was appreciated the least in Germany and the most in America. My argument went like this: The initial Lachmann-Maas procedure involves taking all the witnesses that document the history of a text and dividing them into groups, identifying similarities and differences between those groups. In order to be able to effectively distinguish "genera," Maas introduced the notion of the "indicative error," i.e. an error that is not accidental but indicative of a specific stream of tradition. However, the indicative error concept introduces a subjective element into the reasoning. This label is the result of the researcher's personal decision: It is the scholar who considers a given version to be erroneous, and the error to be indicative. It is hard not to see that with the introduction of the indicative error, the scientific model loses the purity of objective documentation, or even worse: The whole system is threatened even more by the issue of contamination.

Let us assume that we have managed to reconstruct families of manuscripts and each one is a certain ideal type. When it comes to working on the tradition of a specific text, though, things become much more complicated (I encountered such a situation when preparing Cicero's *Pro Milone* for Teubner at the turn of the 1970s and 1980s). An enormous part of the real transmission of ancient texts is impossible to organize in this way. Over several generations, the general belief was that if a researcher did not produce a stemma, it meant he did not know how. In fact, though, traditions often cannot be presented as a stemma of any practical usefulness to an editor because the text's history has not yielded any "genetically pure" models. The ideal model, according to which there primarily exists vertical transmission (younger sources are descendants of older ones and form generation sequences), is a great simplification. Manuscripts coexist

on a horizontal plane as well, and completely ignore chronology in those relationships. People carried codices around with them, copied them, and also sold and bought them without any concern for pure "breeding" of a specific tradition of transmission. Often – usually only along brief sections of a text's history – we can trace the direct descent of certain manuscripts from other, also identifiable codices. However, the world of manuscript sources expanded mainly in a process of constant contamination. A multitude of relationships resulted from the physical contact of individual volumes in circumstances that are impossible (or almost impossible) to reconstruct centuries later. Faced with this ocean of contaminated texts that cannot be grouped into proper families according to the stemmatic model, a classical philologist who believes in Maas is helpless.

Of course for someone whose initiation, like mine, had been based on a tradition as splendidly simplified as *Pro Roscio Comoedo*, the shock was double. In addition, all the traditionally educated classical philologists up to my generation had, since the early 19th century (i.e. from the time of the triumph of German neo-humanism), been schooled to be experts on antiquity (from the 10th century BCE to 400 CE); after that, they were allowed not to be interested in the Middle Ages, only to return as competent experts on Renaissance philology of the 14th and 15th centuries.

Meanwhile, everything that was important for manuscripts happened mainly in the Middle Ages. Manuscripts hid and were copied in isolated communities scattered all across the Christian world of the time. To trace their history, you need to know as much as possible about medieval people and local communities. In-depth knowledge of daily life between the 6th and 13th centuries turns out to be essential. Previously, the author was the editor's partner; however, it appears that now the partner should be the scribe, who traditionally was treated like a pest causing confusion in the sources (Reynolds, 1983).

I began to realize that we should accept this fundamental change of roles. Editors working on a critical edition are more closely related to the scribes whose mistakes and intellectual deficiencies they are used to mocking than to the text's author, M. T. Cicero, whom they worship and whose text they want to restore to its full glory. Thus, in the third year of working on the text of *Pro Milone* I acknowledged the prevalence of contaminated,

"mixed-breed" manuscripts as the main source of doubt regarding the scientific nature of Maas' concept[7].

Not feeling up to becoming a scribe's good partner (for reasons explained in footnote 7), I sought hope in a different area of a textual critic's activity. Because, even if we accept all the accusations against the stemmatic theory as a necessary evil, there emerges an area whose status is even more ambiguous. What does a scholar really do when they move on to *emendatio*? What do I do when I introduce conjecture – amending the text? Don't I make things up, by any chance? After all, I use my able mind to build hypothetical variants of the text that fit my interpretation, my understanding of a given fragment or a whole work. This means I am the one giving the text the shape and meaning that seems the most convincing at the moment of my own reading. Essentially, then, I am practicing literary criticism, and if I change the text in the process, I am involved in artistic and not scientific activity. It's no use pretending that *emendatio* is anything more than a hypothesis, practically impossible to disprove, or that devising it is a form of scientific activity, even just to the extent of *recensio* and *examinatio*.

And so I split my Maas into two parts, agreeing that stemmatics is scientific activity similar to the natural sciences, although tainted with our insufficient knowledge about the way of life and copying of texts in the past, and that it is followed by artistic activity which should be judged according to different rules. I finally wrote a commentary on Maas, which was translated into English, Spanish and Italian, and gave lectures on it in many places, until I realized it was addressed to no one. When artists heard it, e.g. actors and reciters, they liked it very much, but they do not work in editing. Classical philologists, on the other hand, said this division did not apply to them. They insisted that they do very professional work from start to finish, they never make anything up, they have evidence for everything[8].

---

[7] I also started doubting my own skills and competence in textual criticism in this sense. Not only was I no expert on the Middle Ages and had learned paleography as an amateur (despite help from Aleksander Gieysztor); the worst thing was that by personality and disposition, I am a very poor scribe. I make mistakes when copying my own texts, and I do not see those mistakes.

[8] The evidence usually came in the form of arguments showing how a given correction or conjecture could have come from successive corruptions of the text at different stages of transmission, taking into account any hypothetical

3. Present-day methodological debates over textual criticism

Since a new Maas turned out to be a project "for no one," I drew the following conclusions from the experience.

First of all, we need to treat criticism of an ancient text mainly as an important stage in preparing a new interpretation of that text, as a special kind of exercise based on multiple careful readings.

Secondly, during those readings we need to take the entire documentation of the text's transmission and first reduce and simplify it in order to see if the sources can be fitted into a stemmatic model, divided into families, or if it is too discontinuous and contaminated for such a diagram to be helpful to the editor. Later, though, we need to treat the entire set of sources documenting the transmission of the text as witnesses to the text's reading in past times – these are valuable and seldom utilized materials opening up new perspectives of interpretation within the history of a given work's reception (remembering that this documentation encompasses reception from the moment the text was published in antiquity all the way to the 14th/15th century).

Thirdly, the many-generational work of classical philologists, from the Renaissance up to the latest critical editions of a text, should also be considered as a set of resources for the history of reception. In a sense, this is a continuation of the transmission of that text. This is not the place to expound on my own papers written or presented from this position. For the present purpose, it is enough that I have explained why I gave up further publication of model critical editions of Cicero's texts, the kind that Bibliotheca Teubneriana releases. I am convinced that the future of *ars critica* will require a completely different organization and presentation of gathered data, and a different way of combining them with our own interpretations of the world represented in texts as well as documenting the life of those texts over time.

---

manuscripts that were not preserved or were lost. Such reconstructions consist mainly in assuming the occurrence of sequences of paleographic errors. I know how to do that, but I have never viewed this as scientific work, but rather as artistic activity. That is also my perception of the ability to compose successive transformations, none confirmed by sources, for the purpose of producing a new text variant that the editor finds intellectually satisfying.

However, I would like to raise one more question I think is very significant. The relations between textual criticism and methodologies applied in biological sciences are old. I mentioned earlier how geology and paleontology inspired Lachmann, and how evolutionism influenced the development of stemmatic theory. Some say it was the other way round, that biology took stemmatics from Lachmann (Howe et al., 2001; Bod, 2016; Howe, Connolly and Windram, 2012). Whichever it was, evolutionism contributed to consolidating the authority of Lachmann's proposal, even if philologists really were the ones who encouraged naturalists to draw stemmata. It was obvious to Maas that his book presented Darwinism for humanities scholars. In our times, on the other hand, an equally fascinating process is taking place. Researchers of texts are responding to gene theory. We are noticing some amazing similarities. You can look at the transmission of texts like this: Manuscripts, regardless of the people who copy them, behave selfishly and seek their own niches in order to survive. I don't think anyone has called this "manuscripts' will to live," but some have already called it the erotic life of manuscripts.

Sometimes I think manuscripts are like parasites. They seek hosts for themselves. Culture is the environment they live in. In culture, if they want to go on living, they have to find carriers. The hosts/carriers are people who want to read, remember, copy or print these texts. People process them and, passing them on, intentionally or unintentionally change them by placing them in new contexts that form a new environment for them. Because even if people don't change the words, they add new commentaries. Then, the textual sequences of tradition live for as long as they find carriers. Hidden, dormant, they were once able to survive centuries. Today they are much less resilient.

Therefore, we can look at texts as if they were living organisms. This is exactly what came to my mind during my "crisis of faith" in traditional textual criticism, a good generation before such thoughts first appeared in the literature. Unfortunately I kept these bold parallels to myself. But those sentiments led me, even back then in the 1980s, to the conviction that a classical philologist can benefit greatly from the turbulent development of genetic biology. At that time, I had a narrow education, did not have much of an idea about contemporary theory of text, nor was I aware that "the death of the author" had been proclaimed.

Poststructuralism was long alien to me. Meanwhile, I knew quite a lot about biology because I had always been interested in animals and wanted to talk with them. Thus, I found the idea that texts live like species, while manuscripts – the daily focus of my work – live like specimens, very appealing.

Today comparing textology to genetics has become trivial. I find it much more interesting once more to read and think about textual criticism as a practice in which the boundaries between science and the humanities are overstepped to the benefit of research. We keep seeing confirmation that this boundary is arbitrary and that a humanities-trained imagination is useful on the boundary's science side, and vice versa.

In a parallel development, progress in IT is having a revolutionary impact on contemporary textual criticism; it enables enormous databases to be created. This completely changes the perspective of working on "masterpieces" and even more strongly changes the way we look for intertexts (traditionally called similia in classic editions of texts by Latin authors), i.e. borrowings from other works functioning within the work under consideration. Both the critical apparatus and the "similia" apparatus is easy to reduce to an extreme degree, and equally easy to expand immensely. Help from the internet enables a well-prepared group of IT specialists today to envelop every text with much more complete documentation of its genesis than before, and unimaginably more extensive documentation on the presence of the words of others in a given text.

What, then, would be a reasonable compromise between a huge database amassed by computers and the needs of an editor? One answer could be this: After "the death of the author," nothing is forcing us to produce an original (the author's) final text. In such a situation, an editor putting together his critical apparatus and similia should not strive to record any reading of his own; rather, like the author, he should die as an individual (Howard-Hill, 2009; Shillingsburg, 1997, p. 18; Murphy, 2008; Müller-Sievers, 2006).

The intertexts I mentioned before can now be multiplied endlessly, without the editor trying to resolve what and to what extent an author really knew and intentionally imitated. At any time and anywhere, a quotation used by an author without intentional awareness can come alive in individual and collective reception. If we do not want to destroy the picture of the potential audience

network, we have to build multi-tiered similia. We have to be aware that such sets should not be arranged chronologically, because readers' memories do not preserve some of them and ignore others according to chronology. It is the culture in which a text is read at a given time that determines the area of associations.

Referring to my earlier reasoning, I would like to point out that this is another piece of evidence showing that artistic texts are able to find themselves new carriers. For example, a text by X contains formula f which was authored by Y; author X was not aware he was quoting Y, but a reader in another time and place, when Y was popular again, recognized the reference and f became his formula f_1. Thus, the secret life of early texts in different real cultures of reception once again makes good old "influencology" a fascinating way of tracking that secret life (Axer, 1984, 1991).

It seems that textual criticism today can do without humanists. With help from IT technicians alone, computers will build stemmas taking into account all available sources and without using the concept of indicative errors. Computers will also sort the similia as needed, without following any concept of the author's intention and without yielding to any interpretative inclinations of the editor. We will be able to refrain from *evaluatio* and *emendatio*, agreeing that this is not a task for a "scientific editor" but a chance for literary critics to show what they can do[9].

The concept of the death of the editor is being opposed by those who believe that our times require us to openly combine our own interpretation of a work with the form of our critical edition. This means finally moving away from the traditional focus on authors and their intentions toward concentrating on editors and *their* intentions. Instead of committing suicide, editors will have the right to secure their own position in the text (Cybulski, 2014).

The reader of the present paper will easily notice that we have come full circle. Such a position of the editor would justify not only conjecture but also such far-reaching interference

---

[9] A call for such decisions is being loudly voiced by a lobby that sees great opportunities for building new databases and, especially in Poland, is competing for very limited funding against editors who do not hesitate to interfere with the material they are studying. See Bem (2016), a text on why Polish scientific editing does not exist. Of course this author speaks on behalf of editors of contemporary literary texts, but there is no doubt that limiting scientific editing to digital editing is more of a PR slogan than the result of any deeper reflection on the relationship between a canonic text, the author, and the editor. Cf. also Campagno (2012).

as to be reminiscent of reconstructing a nonexistent whole from fragments and fractions[10]. This is not an accusation against editors who see exactly this as being within the scope of their rights – we have simply returned to the fundamental dilemma from which I started this paper. *Ars critica* consists of scientific activity, artistic work and an exercise in forgery of old masters, all rolled into one. Perhaps that is what makes it so fascinating and allows it to bring so many triumphs and disappointments, depending on the temperament, character and mood of whoever is practicing it. As a warning, in the Annex I have provided some information about a great scholar who took the editor's freedom to the extreme.

Annex

I am reminded of something I have always felt to be ethically stirring and intellectually hard to explain. Let me briefly recall a great philologist whose actions brought the condemnation of his peers and posterity upon him, and whom I see as a very early example of an editor assuming the role of an author while fully believing that this serves to preserve priceless cultural heritage which no one else would be able to save. In 1584 in Bologna, the famous philologist and author of the first edition of *Fragmenta Ciceronis* (Venice, 1563), Carlo Sigonio, printed the complete text – lost since antiquity – of the *Consolatio* that Cicero wrote after the death of his daughter Tullia. Only a few fragments and self-quotations scattered across Cicero's writings survived from the piece. Sigonius gave his work the form and poetics of a scientific edition. Since he had hinted earlier that he had access to the manuscript and a heated scholarly debate had erupted around the discovery, opinions for and against were included in the edition. Its success was immediate (six editions within a few years), then disaster struck just as fast. The European elites continued to argue over the piece's authenticity, and Sigonius turned for help to his Polish friends who had earlier been his favorite pupils, including Jan Zamoyski, Andrzej Patrycy Nidecki, and maybe even Jan Kochanowski, but to no avail. There is no doubt today that the old master never came upon an old manuscript that he

---

[10] Actually, Kumaniecki took great pleasure in reconstructing the "lost orations" of Cicero.

proceeded to publish, but reconstructed the whole piece from the preserved fragments, doing so with great expertise. But the verdict was: "He forged it," and Sigonius ended up in textbook studies on forgery in science[11]. In my opinion, Grafton's verdict in "the people vs. Sigonius" is unfair. In this circumstantial case, we need to admit the possibility that the great scholar tried to overstep the boundaries of textual criticism that were acceptable in the Renaissance, proposing a reconstruction hypothesis (like, for example, Grotius when he published Cicero's *Aratea* a generation later). Polish readers should also remember that the same crumbs of the *Consolatio* that Sigonius used to reconstruct his edition were used as material by Jan Kochanowski at about the same time, when he was writing his *Threnodies*. It was just unfortunate that the old professor made an elaborate attempt to conceal what was basically his artistic vision behind evasive explanations referring to a lost manuscript and its unknown finders.

References

Axer, J. (1976). *Mowa Cycerona w obronie aktora komediowego Roscjusza: studium z krytyki tekstu* [Cicero's Oration in Defense of Comedy Actor Roscius: A Study in Textual Criticism]. Wrocław: Ossolineum.

Axer, J. (Ed.) (1976). *M. Tulli Ciceronis Scripta quae Manserunt Omnia* Fasc. 9 *Oratio pro Q. Roscio Comoedo*. Leipzig: B.G. Teubner Verlagsgesellschaft.

Axer, J. (1979). *The Style and the Composition of Cicero's Speech "Pro Q. Roscio Comoedo": Origin and Function* (*Studia Antiqua 3*). Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Axer, J. (1984). Tradycja klasyczna w polskojęzycznej poezji renesansowej a mechanizmy odbioru tej poezji [The Classical Tradition in Polish-Language Renaissance Poetry and the Mechanisms of This Poetry's Reception]. *Pamiętnik Literacki* LXXV, 2, 207–216.

Axer, J. (1989). Kazimierz Kumaniecki – Edytor [Kazimierz Kumaniecki – Editor]. *Przegląd Humanistyczny*, 3, 153–157.

Axer, J. (1991). *Problemy edytorstwa tekstów łacińskich w literaturze polskiej XVI i XVII w.* [Problems of Editing Latin Texts in Polish Literature of the 16th and 17th Centuries]. In: J. Pelc and P. Buchwald-Pelcowa (Eds.), *Problemy edytorskie literatur słowiańskich* [Editing Problems of Slavic Literatures], vol. 1 (pp. 133–148). Wrocław: Ossolineum.

---

[11] According to Anthony Grafton, it was a piece unworthy of Cicero, and it is an inexplicable mystery why the great scholar stood by his forgery; the human soul has its dark side, Grafton also commented (Grafton, 1990).

Bem, P. (2016). *Dlaczego polskie edytorstwo naukowe nie istnieje* [Why Polish scientific editing does not exist]. *Teksty Drugie*, 1, 151–168.

Bod, R. (2016). *How Interactions between Humanities and Science Have Shaped our Knowledge*, Academy Lecture at the Norwegian Academy of Science and Letters, delivered on March 8.

Campagno, D. (2012). Theories of Authorship and Intention in the Twentieth Century. An Overview. *Journal of Early Modern Studies*, *1*(1), 37–53.

Cybulski, Ł. (2014). Krytyka tekstu i teoria dzieła. Jerome McGann wobec anglo-amerykańskiej tradycji edytorstwa naukowego [Textual criticism and the theory of a work: Jerome McGann and the Anglo-American tradition of scientific editing]. *Teksty Drugie*, *2*, 21–48.

Grafton, A. (1990). *Forgers and Critics*, Princeton: Princeton University Press.

Howard-Hill, T. H. (2009). The dangers of editing or the death of the editor. *Ecdotica*, *6*, 284–301.

Howe, C. J., Barbrook, A. C., Spencer, M., Robinson, P., Bordalejo, B., & Mooney, L. R. (2001). Manuscript evolution. *Endeavour, 25*(3), 121–126.

Howe, C. J., Connolly, R., & Windram, H. F. (2012). Responding to criticisms of phylogenetic methods in stemmatology. *Studies in English Literature 1500–1900*, *52*(1), 51–67.

Maas, P. (1927). Textkritik. In: A. Gercke, E. Norden (Eds.), *Einleitung in die Altertumswissenschaft*, part VII, Vol. I (pp. 1–18). Leipzig: B. G. Teubner.

Maas, P. (1958). *Textual Criticism*. Oxford: Oxford University Press.

Müller-Sievers, H. (2006). Reading without interpreting: German textual criticism and the case of Georg Büchner. *Modern Philology*, *103*(4), 498–518.

Murphy, J.S. (2008). The death of the editor. *Essays in Criticism*, *58*(4), 289–310.

Reynolds, L.D. (1983). *Texts and Transmission: A Survey of the Latin Classics*. Oxford: Oxford University Press (reprinted).

Reynolds, L. D., Wilson, N. G. (1968, latest edition 2013). *Scribes and Scholars: A Guide to the Transmission of Greek and Latin Literature*, 4th edition. Oxford: Oxford University Press.

Shillingsburg, P. (1997). *Resisting Texts. Authority and Submission in Constructions of Meaning*. Ann Arbor: University of Michigan Press.

West, M. L. (1973). *Textual Criticism and Editorial Technique*. Stuttgart: B. G. Teubner.

Przemysław Piwowarczyk
ORCID 0000-0002-5812-6672
Faculty of the Humanities, University of Silesia in Katowice

Mechanism of Mystification and Demystification at the Point of Contact between the Humanities and Science: Case Study of the *Gospel of Jesus' Wife*

Abstract

The case of the so-called *Gospel of Jesus' Wife* lays bare all the problems of research on early Christian manuscripts. We have here a manuscript of unknown provenance, a private collector wishing to protect his anonymity, an academic institution hungry for media interest, and, last but not least, the ideological bias of a scholar. In the end, the manuscript proved to be a modern forgery. However, we know this without any serious doubt not thanks to laboratory analyses and more traditional paleographic or historical studies. The issue was not resolved until a professional journalist conducted a journalistic investigation. If this method of verification of manuscript authenticity joins the others, it would be an undisputable benefit of this whole four-year-long saga.

Keywords: Gospel of Jesus' Wife, Karen King, papyrology

Introduction

Almost all the most important manuscripts that bring us knowledge about the religion and culture of late ancient Egypt came to light not in the course of systematic and duly documented excavations but via the antiquarian market, which means uncontrolled digging (sometimes close to plundering) and a chain of agents. Examples include the Nag Hammadi codices (Robinson 2014, pp. 1–119), the Codex Tchacos (Myszor, 2006, pp. 9–12; Nongbri, 2018, pp. 95–96), the Cologne Mani Codex (Koenen, 1973), the Manichaean codices from Medinet Madi (Gardner, Lieu, 1996, p. 148), the library of the White Monastery etc.

When lack of knowledge about the provenance of a manuscript is the standard rather than the exception, forgeries are much easier to procure. The falsification of ancient papyri has its origin at the very dawn of modern papyrology and codicology. We can find biblical papyri forged by Constantinos Simonides (Jones, 2015, 369–373) already in the 1860s, but the problem of forgeries is still one of the most heatedly debated. In November 2018, many scholars published an open letter to Brill publishing house regarding a 2016 volume containing biblical fragments (bought by private collectors in the years 2009–2014) allegedly found in the Dead Sea area. So far, five of them have already been identified as modern forgeries (Mazza, 2018). In the light of such facts, it has become urgently necessary to diagnose the weak points in academia that forgers target. Success depends on the application of procedures that allow effective identification of forgeries. For these challenges, a case study of the so-called *Gospel of Jesus' Wife* (*GJW*) seems particularly instructive.

Outline of the saga

In July 2010, a private collector who wanted to protect his identity contacted Karen King, a professor at Harvard Divinity School, and informed her about a papyrus in his possession which mentions the married status of Jesus. In 2011, he renewed his proposal with a suggestion that he had already found a willing buyer but preferred to give the manuscript to an academic institution. The owner never claimed the manuscript was authentic. He even asked for its authenticity to be verified. He said he had purchased a lot comprising six Coptic papyri in November 1999 from Hans-Ulrich Laukamp, who reportedly bought them in Potsdam in 1963. The proof of the transaction was a photocopy of a sale agreement. The owner also gave King further documentation:

– a copy of a letter from H.-U. Laukamp to Peter Munro from 1982. Munro, an Egyptologist himself, relates in the letter his consultation with Gebhard Fecht who had recognized a fragment of the Gospel according to John among the papyri;

– an undated note by Munro who, once again referring to Fecht, mentions a piece of papyrus containing the words said by Jesus: 'my wife';

– a graphic file with an interlinear translation of the Coptic text into English (presented in public in August 2015).

Before making any public announcement about the papyrus, King consulted with some of the most eminent scholars on papyrology and the Coptic language: Roger Bagnall, AnneMarie Luijendijk, and Ariel Shisha-Halevy.

On September 18, 2012, during the International Congress of Coptic Studies held in Rome, King delivered a paper under the title "A New Coptic Gospel Fragment" which, although it might spark interest among specialists, did not suggest any wider attention outside that circle. The scrap of papyrus has no title, but King labelled it provocatively as the *Gospel of Jesus' Wife*. At the time, a draft of her paper was published on the Harvard Divinity School website. The final text did not appear in the conference proceedings but was published in the *Harvard Theological Review* in 2014 (King, 2014a).

Before the discovery was announced to academics at the congress, Harvard Divinity School informed the media: *The New York Times*, *The Boston Globe* and *Smithsonian* magazine. The Smithsonian Channel planned to produce a one-hour-long documentary about the papyrus with the intention of broadcasting it on September 30. Since severe doubts about the authenticity of the manuscript appeared, the broadcast was cancelled, but work on it had definitely started before September 18. Ariel Sabar, a freelance journalist, interviewed King already two weeks ahead of the Roman congress (Lied, 2016, p. 5); the interview was published on the *Smithsonian* magazine website already on September 17. The day after the official presentation of the discovery, a special press conference was organized in Rome. From the very beginning, the press tried to interpret King's paper in line with political implications, suggesting that Rome was not chosen for the place of the announcement by mere chance (Sabar, 2012b). Such active involvement of the media, exploring new and unexpected aspects of the case, enabled Liv Ingeborg Lied to call the whole process a "saga" (Lied, 2016, pp. 2–3).

Mechanism of mystification

Following the story of the *Gospel of Jesus' Wife*, we can clearly recognize that a skilled forger is not enough to ensure an effective mystification. At least three elements are required:

1. A forger. Now revealed as the owner of the papyrus, Walter Fritz never confessed to fabricating it, and his case was never put before a court, but abundant circumstantial evidence gathered in the course of a journalistic investigation by Ariel Sabar (2016) is very convincing. Fritz graduated in Egyptology in Germany and was a PhD student, so he knew enough to produce a successful forgery. Although his possible motives remain unclear, he had a few good reasons to launch a mystification (financial profits, revenge for personal failures, popularization of his ideas; for each of them there is enough support in the documentation provided by Sabar; see also Depuydt, 2014, 176–177).

2. A scholar. Karen King was deeply engaged in feminist interpretation of early Christianity. She highlighted the vital role of women in the early decades of the new religion, later relegated to the margins by patriarchal clericalism (King, 1988; King, 1997; King, 2003). She was by no means isolated in her interpretations, despite the fact that part of public opinion regarded such views as ideologically biased: "Mary Magdalene has become a project for a certain kind of ideologically committed feminist scholarship" (Woodward, 2003). Already the first reactions to her paper noted that the papyrus as a potential forgery might be a deliberate attempt to compromise the feminist approach.

3. An institution. Media coverage of the discovery and the creation of an 'event' restricted to selected media as a piece of exclusive news would be virtually impossible without the involvement of an academic institution like Harvard Divinity School. King's role in this process remains unknown. This dynamic between academia and the media resulted in a strategy of promotion that had the label "the Gospel of Jesus' wife" at its center – a title without a doubt chosen with full awareness of its reception in the media.

Mechanisms of demystification

From the very moment of the public presentation of the papyrus, its authenticity was put into question. We can point to three approaches taken to verify or negate the genuineness of the manuscript.

1. Laboratory analyses

The papyrus was subjected to various laboratory tests and could be counted among the scientifically best analyzed manuscripts

in history. The decision of Harvard Divinity School deserves appreciation, because the owners of ancient manuscripts (including institutions) often do not allow destructive analyses (such as radiocarbon dating) for fear of the early dating of objects in their possession being challenged (Nongbri, 2018, 270). On the other hand, however, the official website devoted to the Gospel (https://gospelofjesuswife.hds.harvard.edu/) contains no other results besides these scientific analyses, as if laboratory tests alone could resolve the issue once and for all.

a. Radiocarbon analysis resulted in dates ranging from the 7th to the 9th century AD (Tuross, 2014; Cf. King, 2014, p. 135)

b. Raman analysis proved that the ink used on the papyrus does not differ from ancient ink recipes (King, 2014, pp. 134–135)

c. Fourier-transform infrared microspectroscopy proved the homogeneity of the chemical composition of the papyrus in all its parts and showed that the patterns of oxidation are compatible with ancient manuscripts.

d. Scrutiny of the damaged area of the papyrus did not prove the presence of ink on the lower layer of the fibers, which would indicate that the ink was put on previously damaged material (Choat, 2014, p. 160–161).

As long as the discussion on the manuscript's authenticity only centered on scientific methods, no definitive conclusion could be reached. Although there was no proof against the physical authenticity of the papyrus, nevertheless forgery of the texts was not excluded. We need to bear in mind that a skilled forger is fully aware of standard scientific procedures (and we indeed know that Fritz used to apply infrared to read pharaonic papyri, Sabar, 2016), and that even an amateur can buy an ancient scrap of blank papyrus online and produce an imitation of an ancient ink at home. Indeed, it turns out to be relatively easy to fabricate a fake that would positively pass laboratory examination. In such a case, the cult of hard science would make demystification even more difficult.

2. "Traditional" methods of verification

Already in September 2012, a draft of King's paper appeared online. Also, another scrap bought together with the *GJW*, containing a fragment of the Gospel according to John, was also publicized and relatively quickly identified as a forgery (Bąk,

2016b). The text of the Gospel was immediately caught in a heavy crossfire of criticism. Interestingly, a detailed discussion took place not only (even not predominantly) in peer-review journals but on scholarly blogs.

Paleographic analyses did not unanimously prove whether this was a modern forgery. When some scholars pointed to the untypical ductus of the writing and the application of a brush instead of a pen (kalamos) (Lundhaug, Suciu, 2012), others gave examples of authentic papyri featuring the same specification (Choat, 2014). In the case of paleography, the lack of definitive conclusions is indeed the rule rather than the exception.

The accumulation of meaningful phrases on such a tiny and mutilated scrap of papyrus also fell under suspicion. Hugo Lundhaug and Alin Suciu write that "we welcome anyone to try to cut out a piece of this size from any literary codex from late antiquity and get a result that is as easy as this one to make sense of and interpret" (Lundhaug, Suciu, 2012). But also in this case, other scholars disagree, accentuating the difficulty of putting the separate phrases together into coherent passages that would fit into the limited space of the lost parts of the papyrus (Watson, 2012b).

Textual relations between the *GJW* and the *Gospel of Thomas* (preserved in the second codex from Nag Hammadi) were evident to scholars from the beginning of the discussion. Nevertheless, they differ in the conclusions drawn from that fact. For some, the *Gospel of Thomas* was the source of textual excerpts used to fabricate a patchwork (Depuydt, 2014; Bąk, 2016a, pp. 74–82), for King the dependence was only on a literary level – nothing unexpected in the Christian Apocrypha (King, 2014a, p. 157; King, 2014b, p. 193).

More intriguing was the fact that the pattern of the lines of text in the *Gospel of Jesus' Wife* is the same as in the Nag Hammadi manuscript (and editions) of the *Gospel of Thomas* (Watson, 2012a). Especially striking is the first line, which begins and ends in almost perfect accordance with the divisions in the Nag Hammadi codex (NHC II 49,36 – 50,1 = *GosThom* log. 101,1–2, Gathercole, 2015, 304–305). Besides this, the *GJW* has many orthographic and syntactical peculiarities that cannot be explained by purely literary interdependence between the two texts. The only possible solution was the direct dependence of King's manuscript on the Nag Hammadi codex, which appears

extremely unlikely, especially since we would have to assume a very faithful tradition of the *GJW* text throughout the four centuries after the Nag Hammadi codex was buried in the 4th (or 5th) century (Gathercole, 2015, 302–312).

It was even possible to trace that the forger used an interlinear translation by Mike Grondin published online as a pdf file in 2002. The *Gospel of Jesus' Wife* duplicates a typographical error making the Coptic text grammatically incorrect (Gathercole, 2015, 304–305). Similarly, the English translation provided by the owner of the papyrus comes from the same edition. It is reasonable to think that the forger reassembled the phrases isolated from their original context in the Coptic as well as in the English text (Bernhard, 2017). Although some errors occur in the translation, they might have been made deliberately to suggest a lack of language expertise (according to the owner, the translation was made by a Coptic priest).

Besides many doubts raised by the text itself, King's interpretation of the fragmentarily preserved lines was based on a formula unattested in any other early Christian text. Undeniably, some Christians were interested in the special role of Mary Magdalene in the circle of Jesus' disciples, some texts even testified that she was thought to be initiated in a special way into his salvific mission. However, not a single text recognizes Mary Magdalene as Jesus' wife; not one even suggests that Jesus was married at all (Myszor, 2013, 170–172; Gathercole, 2015, 294–302). Meanwhile, this is the path King's interpretation follows, and it is no wonder that far-reaching speculation exploded in some media and among a wider audience (Sabar, 2012b), though King never suggested that the papyrus could contribute something to our knowledge about the historical Jesus.

3. Journalistic investigation

Although the heated discussion among biblical scholars and papyrologists provided solid premises for forgery, arguments based on the laboratory tests, as being founded on hard science, were put forward all the time against evaluation based on paleography (which actually was not unequivocal), content and historical probability. The involvement of a professional journalist was crucial for breaking this stalemate. Ariel Sabar had covered the saga from its start and, when the suspicion of forgery was

raised, conducted a thorough investigation in the United States and Germany (Sabar, 2016). There is no place or need here to summarize the whole story. His article in *The Atlantic* is worth reading not only because of the facts but also for its admirable journalistic methodology. The reconstructed history behind the *GJW* turns out to be exceptionally sensational. Sabar used many methods to obtain information. He combined online resources (starting from web browsing, through searching publicly available databases on persons and companies, to research on porn fora, which also yielded some important pieces of the puzzle), personal interviews (with members of families, former employees and familiars of persons mentioned in the documents given to King and others deduced from them), verification of the authenticity of modern documentation, and even reading *The Da Vinci Code* (this – it appears – might have been the forger's inspiration). None of these methods belongs to the repertoire of a biblical scholar or papyrologist.

Conclusions

Some more general conclusions can be drawn from the whole discussion over the authenticity of the *GJW*. On the one hand, it is very clear now that neither traditional judgements based on the experience and intuition of scholars nor laboratory tests (which is especially worth underlining) are able to provide definitive arguments when we are dealing with a tangle of personal, institutional and media business interests. This is especially the case when the truth about a given artefact lies not only in itself but also in the complicated biographies and shady transactions that contributed to its creation. It seems that the space for journalists in the process of discovering the past has become much wider than before. In September 2016, at a conference focused on manuscript forgeries, besides scholars the speakers were also journalists participating on equal terms: Ariel Sabar and Nina Burleigh who worked on another infamous fake of recent years, the so-called ossuary of James, brother of Jesus (Mazza, 2016).

On the other hand, academia should be more cautious when collaborating with the media, especially in profiling its message to fit promotional strategies. All in all, the whole affair with the *GJW* was the effect of scholarly debate being obstructed by media involvement.

References

Bąk, T. (2016a). „Jezus rzekł do nich: Moja żona". Znaczenie kontrowersyj-nego manuskryptu koptyjskiego w świetle współczesnej wiedzy i polemiki ["Jesus said to them: My wife." The significance of the controversial Coptic manuscript in the light of contemporary knowledge]. *Collectanea Theologica*, *86*, 61–83.

Bąk, T. (2016b). Harvard Lycopolitan John - kolejne fałszerstwo? Koptyjski manuskrypt w świetle Ewangelii Żony Jezusa i współczesnych badań [Harvard Lycopolitan John: Another forgery? A Coptic manuscript in the light of the Gospel of Jesus' wife and its contemporary examinations]. *The Biblical Annals*, 6, 227–249.

Bernhard, A. (2017). Postscript: A final note about the origin of the Gospel of Jesus' wife. *New Testament Studies*, *63*, 305–317.

Choat, M. (2014) The Gospel of Jesus's wife: A preliminary paleographical assessment. *Harvard Theological Review*, *107*, 160–162.

Depuydt, L. (2014). The alleged Gospel of Jesus's wife: Assessment and evalu-ation of authenticity. *Harvard Theological Review*, *107*, 172–189.

Gardner, I., Lieu, S. (1996). From Narmouthis (Medinet Madi) to Kellis (Ismant El-Kharab): Manichaean documents from Roman Egypt. *Journal of Roman Studies*, *86*, 146–169.

Gathercole, S. (2015). The Gospel of Jesus' wife: Constructing a context. *New Testament Studies*, *61*, 292–313.

Jones, Ch. (2015). The Jesus' wife papyrus in the history of forgery. *New Testament Studies*, *61*, 368–378.

King, K. (1988). *Images of the Feminine in Gnosticism*. Philadelphia: Fortress Press.

King, K. (1997). *Women and Goddess Traditions: In Antiquity and Today*. Minneapolis: Fortress Press.

King, K. (2003). *The Gospel of Mary of Magdala: Jesus and the First Woman Apostle*. Santa Rosa: Polebridge.

King, K. (2014a). "Jesus said to them, 'My wife…'": A new Coptic papyrus fragment. *Harvard Theological Review*, *107*, 131–159.

King, K. (2014b). Response to Leo Depuydt, "The alleged Gospel of Jesus's wife: Assessment and evaluation of authenticity". *Harvard Theological Review*, *107*, 190–193.

Koenen, L. (1973). Zur Herkunft des Kölner Mani-Codex. *Zeitschrift für Papyrologie und Epigraphik*, *11*, 240–241.

Lied, L.I. (2016). Media Dynamics and Academic Knowledge Production: Trac-ing the Role of the Media in the Gospel of Jesus' Wife Saga. Retrieved from http://www.academia.edu/28624547/Media_Dynamics_and_Aca-demic_Knowledge_Production_Tracing_the_Role_of_the_Media_in_the_Gospel_of_Jesuss_Wife_Saga.

Lundhaug, H., Suciu, A. (2012). On the So-Called Gospel of Jesus's Wife. Some Preliminary Thoughts. Retrieved from https://alinsuciu.com/2012/09/26/on-the-so-called-gospel-of-jesuss-wife-some-preliminary-thoughts-by-hugo-lundhaug-and-alin-suciu/.

Mazza, R. (2016). Fragments of an Unbelievable Past? Constructions of Provenance, Narratives of Forgery. A Report. Retrieved from https://facesandvoices.wordpress.com/2016/09/19/fragments-of-an-unbelievable-past-constructions-of-provenance-narratives-of-forgery-a-report/.

Mazza, R. (2018). *Open letter to Brill: Fake and unprovenanced manuscripts*. Retrieved from https://facesandvoices.wordpress.com/2018/11/05/open-letter-to-brill-fake-and-unprovenanced-manuscripts/.

Myszor, W. (2006). *Ewangelia Judasza* [The Gospel of Judas]. Katowice: Księgarnia Św. Jacka.

Myszor, W. (2013). Koptyjska „żona Jezusa" [The Coptic "wife of Jesus"]. *Śląskie Studia Historyczno-Teologiczne*, *46*, 168–172.

Nongbri, B. (2018). *God's Library. The Archaeology of the Earliest Christian Manuscripts*. New Haven: Yale University Press.

Robinson, J.M. (2014). *The Nag Hammadi Story*, vol. 1, *The Discovery and Monopoly*. Leiden: Brill.

Sabar, A. (2012a). The Inside Story of a Controversial New Text About Jesus. Retrieved from https://www.smithsonianmag.com/history/the-inside-story-of-a-controversial-new-text-about-jesus-41078791/.

Sabar, A. (2012b). Update: The Reaction to Karen King's Gospel Discovery. Retrieved from https://www.smithsonianmag.com/history/update-the-reaction-to-karen-kings-gospel-discovery-84250942/.

Sabar, A. (2016). The Unbelievable Tale of Jesus's Wife. Retrieved from https://www.theatlantic.com/magazine/archive/2016/07/the-unbelievable-tale-of-jesus-wife/485573/.

Tuross, N. (2014). Accelerated mass spectrometry radiocarbon determination of papyrus samples. *Harvard Theological Review*, *107*, 170–171.

Watson, F. (2012a). The Gospel of Jesus' Wife: How a Fake Gospel-Fragment Was Composed. Introduction and Summary. Retrieved from http://ntweblog.blogspot.com/2012/09/the-gospel-of-jesus-wife-how-fake.html.

Watson, F. (2012b). Addendum: The End of the Line? Retrieved from http://ntweblog.blogspot.com/2012/09/francis-watson-addendum-end-of-line.html.

Woodward, K. (2003). A Quite Contrary Mary. Like Jesus, Mary Magdalene is now the subject of a cultural makeover. What agenda do feminist scholars have in mind? Retrieved from https://www.beliefnet.com/entertainment/movies/the-da-vinci-code/a-quite-contrary-mary.aspx.

Karol Wilczyński

ORCID 0000-0002-9847-2924

Jagiellonian University

»Artes Liberales« Academy

Why Is Philosophy Bad for the Soul? Commentary on Al-Ġazālī's Critique of the Philosophers[1]

Abstract

In this text, two notions described in two works by Al-Ġazālī are analyzed: the notion of "precipitance" (Arabic: *tahāfut*) and "attachment to authority" (Arabic: *taqlīd*), which are described in *Tahāfut al-falāsifa* (English: *The Precipitance of the Philosophers*) and *Al-Munqiḏ min al-ḍalāl* (English: *Rescuer from Error*). I try to show that Al-Ġazālī in his criticism of philosophy focused not only on theoretical issues, but that one of the key parts of his criticism are practical issues concerning the philosophers' way of life. Pointing to specific examples of using the two aforementioned concepts to overthrow philosophy, I propose a different interpretation of these two late works by Al-Ġazālī.

Keywords: Al-Ġazālī, Islam, philosophy, erring

Al-Ġazālī, a Muslim thinker who lived at the turn of the 11th and 12th centuries, is known to this day not only as the author of texts on the theology and spirituality of Islam, but also as a sworn opponent of falsafa – the philosophy developed by Arab authors between the 9th and 14th centuries within the Caliphate and other Muslim countries. In two of his works, *Tahāfut al-falāsifa* (*The Precipitance of the Philosophers*, also translated as *The Incoherence of the Philosophers*) and *Al-Munqiḏ min al-ḍalāl* (*Rescuer from Error*), he focuses on overthrowing the arguments and concepts attributed to the authors of falsafa.

In the present text, however, I would like to demonstrate that besides views and theories, in his critique Al-Ġazālī also says a great deal about the life practice of Arab philosophers. Moreover, it was two fundamental errors of action, which Al-Ġazālī calls "precipitance" (*tahāfut*, cf. Treiger, 2012) and "conforming to authority" (*taqlīd*), that – according to the author – led the philosophers to make serious cognitive mistakes. The present paper will analyze Al-Ġazālī's position on these issues. In the first part, I will focus on a brief presentation of the ideological context in which the author of *Rescuer from Error* wrote his works, and then will move on to an analysis of the two ideas he developed: precipitance, and *taqlīd* as stemming from it. Part two will discuss the consequences, which also appear significant from a contemporary perspective, of Al-Ġazālī's critique and the remedial measures he proposes, which he believed would help overcome trends harmful to philosophy. To conclude, I will briefly outline the impact that my proposed interpretation of Al-Ġazālī's works could have on the evaluation of his entire oeuvre in the context of contemporary research (cf. e.g. Garden, 2014; Griffel, 2009; Treiger, 2012).

1. Ideological context

Understanding the sense of Al-Ġazālī's oeuvre also requires understanding the purpose of specific works he wrote in a historical, cultural and ideological context. Al-Ġazālī was (and still is) known mainly as the author of one of the most important works of Muslim religious literature: *The Revival of the Religious Sciences*. In his lifetime he also headed the famous Niẓāmīya madrasa, a Qur'anic school in Baghdad (Günther, 2010, pp. 15–35). His long-time master was the famous thinker Al-Ġuwaynī (d. 1085), who represented the Ashʿarite school of theology. It was to him that Al-Ġazālī owed the possibility to explore various sciences, including Aristotelian and Avicennian philosophy (Garden, 2014, p. 18 and Griffel, 2009, pp. 30–31). It needs remembering, though, that Al-Ġuwaynī primarily taught kalām and Ashʿarite theology. This means falsafa was most likely presented in an openly critical approach.

Taking part in the battles between theological and political groups of his time, Al-Ġazālī rejected Ashʿarism as the proper

method of theological studies. In addition, he involved himself in the struggle for the succession of his patron, the vizier of the Seljuk state, Niẓām al-Mulk, murdered by supporters of Hassan as-Sabbah who were later called the Nizaris or Assassins; this was a faction of the Ismailis to whom Al-Ġazālī devoted several polemic works. It was probably in the context of the fierce battle for political and religious power that accusations compromising Al-Ġazālī as a Muslim cleric appeared, charging him with belonging to the falsafa group. The rulers of the Caliphate at the turn of the 11th and 12th centuries considered philosophy to be a heterodox trend in Islam. Contemporary researchers (e.g. Abul Quasem, 1974; Treiger, 2012) point out that this was the period when *The Precipitance of the Philosophers* and *Rescuer from Error* were written, and they, too, should be interpreted as polemical works. It is important to remember that it is in these two works above all that Al-Ġazālī focuses his criticism also on the life practice of philosophers.

2. Tahāfut

What is the "precipitance" from the title of one of Al-Ġazālī's main works on philosophy, the aforementioned *Tahāfut al-falāsifa* (cf. Marmura, 2000)? In the Latin Middle Ages, due to the title having been translated as *Destructio Philosophorum* ("the destruction of the philosophers"), a myth arose and then functioned for many centuries that Al-Ġazālī was the author of the definitive work refuting the theories of the ancient and medieval philosophers. However, *tahāfut* does not only mean questioning theoretical concepts. In one of his earlier treatises, *Criterion of Action*, Al-Ġazālī points out that "precipitance" is the main mistake made by those who strive for happiness and choose it as the practical goal of their lives. "Precipitance," he claims, "accompanies them incessantly, filling every hour of the day, leaving them unable to identify what desires truly drive them" (*Mīzān al-ʿamal*, p. 25, in: Bīǧū, 2008). In *The Niche for Lights*, meanwhile, in a fragment describing the development of the "power of imagination," Al-Ġazālī writes about a child who:

> cries and demands something even if it is hidden from him, because it appears in his imagination. Some animals have this ability, but others do not. Moths, which are lost when they fly toward firelight

(Arabic: *al-mutahāfit ʿalā l-nār*) do not have [imagination] because they gather around the flame striving for the sun's light, confusing the lamp with an open window leading them to the light. They rush at [fire] which injures them. And if they fly by and find themselves in darkness again, they return to [the light] once more. If they remembered the pain this caused them, they would not return to [that light], having been burned once. So does a dog that has taken a beating with a stick run when he sees it again (*Miškāt al-anwār*, p. 36, in: Buchman, 1998).

What is more, however, to Al-Ġazālī precipitance means the lack of proper mental practice consisting of appropriate inner exercise that serves to cleanse the intellect (as a tool of cognition) of anything that prevents it from focusing on the subject of study. Writing about this cleansing process, Al-Ġazālī indicates that precipitance is the effect of its lack or – as in the case of the philosophers – it being performed wrongly. As he writes in *The Niche for Lights*, the necessity to properly cleanse the intellect stems from the fact that the power of imagination, which is closely linked to it, is like "glass" which, on one hand, "lets through" whatever ultimately reaches the intellect. Like glass, however, the imagination is made "from earthly clay" and therefore requires constant "cleaning" and "polishing." In the metaphors used in the treatise *The Niche for Lights*, reason is like a lamp that nevertheless has to be protected from "strong upheavals" with (well-cleaned) glass, i.e. the power of imagination, and also memory which is related to it. As Al-Ġazālī writes, pure imagination is essential in the process of studying, something the philosophers seem to forget:

> It [i.e. imagination] is very much needed, because it enables rational knowledge to be organized, thanks to which it is not mixed up, shaken and scattered … Images from the power of imagination are very helpful to rational knowledge … It is thanks to cleaning and polishing it like glass that it does not obscure the light of the lamp [i.e. reason], but rather passes it on in an appropriate way. It also protects the light from destructive winds and strong upheavals (*Miškāt al-anwār*, pp. 39–40).

Precipitance is described in a similar way in *Tahāfut al-falāsifa*. Here, it applies mainly to the study practice of philosophers, which to Al-Ġazālī is often connected with excessive faith in everything cognition brings, and with an uncontrollable desire

to learn. Practically every chapter (discussion on a given view) in *The Precipitance of the Philosophers* ends with the author's conclusion that the philosophers put too much trust in the capability of human reason. For example, summarizing the twelfth discussion, on the self-knowledge of God, which according to the author the philosophers are unable to prove, Al-Ġazālī writes:

> That is why, out of necessity, they must reject His knowledge of Himself, since nothing but [His] will justifies the existence [of the self-knowledge of God] and nothing proves it besides the fact that the world came into being in [a specific] time. Interpreting this doctrine [i.e. on the existence of the self-knowledge of God] as false, one can consider everything to be false if one thinks about these things using reason alone … There is no astonishment or surprise in the reflection of these confused minds on the names of God. The only astonishing thing is the arrogance found in their souls (Arabic: *aġibuhum bi-ānfasihin*) and their faith in their own arguments. What surprises is their conviction that they know of these matters with absolute certainty, despite the confusion and errors (Arabic: *al-ḫabṭ wa al-ḫabāl*) they commit (*Tahāfut al-falāsifa*, p. 133).

But what causes errors in the practice of the philosophers in their studies? In Al-Ġazālī's opinion, precipitance is connected with a special attitude toward one's own mind, namely excessive faith in personal cognitive capabilities:

> I have noticed that a certain group of people believe in their own superiority over others, because they are more intelligent and know more. They have abandoned all the religious obligations fulfilled by Muslims. They mock the good recommendations of Islam, above all the practice of piety and avoiding what is forbidden. They disregard what Sharia law teaches. Not only do they ignore bans, but they reject faith in Islam by various speculations and following the example of people 'who repel from the path of God, who strive to make it crooked and who disbelieve in the Hereafter' (cf. Qur'an 7:45) (*Tahāfut al-falāsifa*, p. 1).

Therefore precipitance is linked to pride – the suspension of a kind of self-criticism or self-reflection in relation to one's beliefs – which leads a person to lose the right path of cognition and, consequently, life practice. To Al-Ġazālī this means, of course, rejecting Islam. In the summary of the third discussion, Al-Ġazālī writes about the philosophers exploring the mysteries of the hereafter or God's names as being people who:

stray from His straight path and lose their way on the road of His leadership; they also contradict His words: 'I did not take them as witnesses to the creation of the heavens and the earth or to the creation of themselves' (Qur'an 18:51). They think about God with the wrong words, and, finally, they believe that the depth of His divine nature could be captured with the help of human powers. They are full of arrogance with regard to their minds (Arabic: *al-maġrūrīn bi-ʿuqūlihim*) (*Tahāfut al-falāsifa*, p. 71).

In the introduction to *The Precipitance*, meanwhile, Al-Ġazālī indicates that this "arrogance" or "pride" leads the philosophers to think they are better than "the Prophet and His companions," that they no longer have to pray and perform the rituals prescribed by Islam (*Tahāfut al-falāsifa*, p. 1) and thus also lead astray ordinary faithful Muslims (*Tahāfut al-falāsifa*, p. 2). Therefore the ultimate effect of precipitance – understood to mean the erroneous practice and inappropriate working of the human mind – is, in Al-Ġazālī's view, that the philosophers will be unable to achieve knowledge of the truth, and consequently to achieve happiness, because "achieving a goal and salvation are only possible thanks to knowledge and practice [realized] together" (*Mīzān al-ʿamal*, p. 26).

3. Taqlīd

This is not the end of the critique of the philosophers. As Al-Ġazālī notes, precipitance causes them to accept without reflection anything they have taken from the ancient philosophers. Their error, therefore, is both *tahāfut* and *taqlīd*:

> The deviation of these people stems from … mindless acceptance of everything they hear or see around them. They could not avoid this. They were raised in a non-Muslim environment, and their ancestors followed the wrong path. Secondly, their deviation comes from cognitive errors: skepticism, inappropriate leadership and stupidity based on fabricated ideas … When I saw the emptiness of these confused fools, I decided I would write this book in order to refute the ancient philosophers. It will show their precipitance and the error of their views (*Tahāfut al-falāsifa*, p. 2).

The notion of *taqlīd* is decidedly negative in character in Al-Ġazālī's anti-philosophical writings, and is linked to blind imitation of a given teacher or doctrine, which leads a person to

switch off their own inner mental effort in acquiring knowledge of the truth (but at the same time, it can be the effect thereof, cf. Menn, 2003, p. 148). *Rescuer from Error*, an autobiography, shows the original vision of how Al-Ġazālī freed himself of the *taqlīd* of Ismailite doctrine as well as falsafa and kalām. In a sense, according to Al-Ġazālī *taqlīd* is described as a "natural" attitude: Children of Muslims, but also followers of other religions, acquire their main beliefs at home and thus accept the authority of those religions. But Al-Ġazālī points out the simple fact that this is precisely why everyone leaves home with the religion they were raised in, but these beliefs (Arabic: *iʿtiqād*) are not the same as knowledge (Arabic: *ʿilm*).

The nominal aim of *Rescuer from Error* is to show how Al-Ġazālī achieved a level of reliable knowledge, by rejecting the senses as its source on one hand, and on the other – an important consideration – rejecting the teachings of various religious groups and the teachings of influential philosophers (*Al-Munqiḏ min al-ḍalāl*, p. 15, in: Jabre, 1969). Al-Ġazālī points out, however, that contrary to the Caliphate's other intellectual traditions, a blind attachment to falsafa leads not only to intellectual laziness or erroneous conclusions from one's studies, but also precipitance and pride. This is, according to the author, related particularly to studies in logic and mathematics, which at the time were considered an essential introduction to philosophy. In Al-Ġazālī's view, however, young students often fell into an unusual trap:

> the first flaw [of philosophy] is visible to anyone studying mathematics, namely the amazing precision and clarity of proofs. That is why he begins to trust the philosophers more, and to think that all sciences resemble mathematics in terms of clarity and reliability. … Paradigms in mathematical sciences are categorical in nature, whereas in metaphysics they are probable (Arabic: *taḥmīn*). But if someone follows *taqlīd*, they will not accept this [aforementioned conclusion] but, guided by a desire for, an absurd[2] drive toward and a love of [their own way] of reasoning, they will insist on a good opinion of all sciences [and thus mistakenly insist on the categorical nature of all sciences]. And this is a huge mistake that should be a warning

---

[2] The ambiguous term *bāṭil* is analyzed, among others, by Linant de Belle-fonds, Y. (2012), "Fāsid wa Bāṭil", in: P. Bearman et al. (Ed.), *Encyclopaedia of Islam. Second Edition*, retrieved from http://dx.doi.org/10.1163/1573–3912_islam_COM_0215 and Rosenthal, F. (2014), p. 420.

to everyone who begins studying these [mathematical] sciences. Even if they are not related to the issue of religion, because they are part of the foundation of their sciences, their wrongdoing and corruption [i.e. that of the philosophers who follow *taqlīd* and amend sciences in absurd ways] will also infect those studying [mathematical sciences]. Few [people studying mathematics] are not stripped of religion, willingly taking the bridle of piety off their heads (*Al-Munqiḏ min al-ḍalāl*, p. 21)[3].

Al-Ġazālī shows that pride – the most serious accusation against the practice of the philosophers – already applies to beginner students, not just their masters. Enchanted by the clarity and precision of mathematical reasonings, they conclude that the same properties apply to metaphysical reasonings, including those that consider religious issues. What is important, however, is that following this mistaken belief which gives them an illusory sense of certainty, they take off the "bridle of piety," namely the tool thanks to which people can be guided by God and true religion in their life practice. Pride connected with wrongful practice and knowledge unaware of its own limitations ultimately leads to unbelief. In other words, philosophy threatens to lead people to abandon Islam, which Al-Ġazālī considers to be tantamount not only to unhappiness on earth but also to eternal damnation. Such a threat is due to the fact that the philosophers, in the author's view, have mixed truth (adopted from the prophets) with falsehood (cf. *Al-Munqiḏ min al-ḍalāl*, pp. 25–26). What is worse, however, the attachment to reason is so great that students of philosophy fall into the trap of *taqlīd*, thoughtless imitation that eliminates any training of the imagination. Hence, as Al-Ġazālī writes in *The Precipitance*, we need to see that as far as knowledge coming from falsafa is concerned, "ignorance is closer to salvation than reasoning stripped of faith. Blindness is closer to the whole than cross-eyed vision" (*Tahāfut al-falāsifa*, p. 3).

---

[3] "Fa-haḏā īḏa qarrara ʿala haḏā al-alaḏī alḥada bi-taqlīd lam yaqaʿu minhu mawqiʿ al-qubūl bal taḥmaluhu ġalba al-hawa wa al-šahwa al-bāṭila wa ḥub al-takāīs ʿala ān yuṣarru ʿala taḥsīn al-ẓann bi-him fī al-ʿulūm kulluhā. Fa-haḏihi āfa ʿaẓīma li-āġiliha yaġibu zaġr kul man yaḥūḍu fī tilka al-ʿulūm fa-ānnahā wa īn lam tataʿallaqu bi-ʾāmr ad-dīn, wa lakin lamā kānat min al-mabādīʾ ʿulūmihim saraa īlayhi šarihim wa šuʿumihim fa-qalla man yaḥūḍu fīhā illā wa yanḥalaʿu min ad-dīn wa yanḥalu ʿan rʾasihi li-ǧām al-taqwa." I would like to thank Anna Wilczyńska and Tomasz Pietrzak for their help with the translation of this obscure excerpt. Comments are added in square brackets – K.W.

In the end, then, an improper relationship between the lower and higher powers of the human mind seems to be the source of the philosophers' problems. In the introduction to *The Revival of the Religious Sciences*, Al-Ġazālī criticizes "false scholars" who "cause people to imagine that knowledge means no more than a decree of someone in power (Arabic: *fatwā*) … or a theological argument in a discussion (Arabic: *ǧadal*) in which everyone, guided by vanity, sharpens their weapons to defeat their opponent … or a poem (Arabic: *saǧʿ*) thanks to which a preacher wins the people's acclaim" (*Iḥyāʾ ʿulūm al-dīn*, vol. 1, p. 2). In the context of the above, it appears that an improper attitude to knowledge leading to precipitance means that one should also count the philosophers among "false scholars." Al-Ġazālī definitely had a low opinion of them. On many occasions, he portrays them as authors who strive for knowledge solely because of "earthly" things, those that concern themselves, their careers and accolades above all. Thus, what Al-Ġazālī condemns is not knowledge itself, but a purely worldly, mercenary and ambition-driven approach to it. In his view, this is the fundamental mistake of the philosophers (cf. Kukkonen 2016). That is the reason why Al-Ġazālī takes it upon himself to "undo the knots of [false] opinions," which is meant to lead to liberation from "blind imitation of authority" (*Al-Munqiḏ min al-ḍalāl*, p. 15). Incidentally, it worth adding that Al-Ġazālī judged his role to be a commendable one, as that of someone who brings new life:

> Ignorance is the worst death, and knowledge is the best life. God, glory be to Him, mentioned knowledge and ignorance in His holy book and called them life and death (cf. Qur'an 51:10–13). Whoever lifts anyone from ignorance to knowledge creates a new man and gives him a new, blessed life (*Al-Maqṣad al-asnā*, p. 136, in: Shehadi, 1971).

4. Counteracting precipitance

As a "savior from error" bringing "rescue from the wrong path" of the philosophers, Al-Ġazālī devoted a lot of attention to ways of counteracting precipitance of the mind and an improper attitude toward authorities. The main tool Al-Ġazālī indicates as being effective in the proper shaping of one's mind is the aforementioned "cleansing of the soul." A practice often mentioned by the author of *The Precipitance of the Philosophers* is reminding

oneself of one's intellectual capabilities, exercising the imagination that connects the soul to the body, enables one to remember about the need to have "both feet firmly on the ground." Secondly, it is the imagination – as a power to which memory is linked – that enables one to recall mistakes made in the past; in other words, it stimulates self-reflection. Importantly, Al-Ġazālī believes the cleansing process, although necessarily performed by oneself, can be made more efficient by "a guide along a wise path" (Arabic: *wāʿiḏ zakiyy al-sīra, Mīzān al-ʿamal*, p. 25). This can be accomplished, for example, by stimulating "appropriate images in the imagination" so that "doubt" is born in the mind. As Al-Ġazālī admits, the "pride" of the philosophers is so great that it is impossible for anything to "amaze" them in their state of mind. They are already "intellectually barren" (*Tahāfut al-falā-sifa*, p. 106), therefore they are unable to stimulate their own minds – they need someone (or possibly something, e.g. the right reading matter) to do it for them (cf. Garden, 2014, p. 56). That is why Al-Ġazālī casts himself in the role of a "snake charmer" capable of handling the snake (symbolizing falsafa: cf. *Al-Munqiḏ min al-ḍalāl*, p. 28). He can thus extract what is valuable from it, and even produce an antidote to the snake's venom; similarly, taking the writings of the philosophers, he will be able to draw from them whatever is of value to society and make it immune to their errors.

Importantly, Al-Ġazālī displayed a similar attitude toward other intellectual traditions of his times, including Sufism which, according to many popular studies, he supposedly represented. In the treatise *Criterion of Action* (*Mīzān al-ʿamal*), Al-Ġazālī describes a state of excellence as the appropriate combination and use of two life paths: that of theorists (i.e. the philosophers) and that of Sufis, while pointing also to their insufficiency. What is important, he makes use of an unusual literary form for this: a tale about a ruler. The king in this story invited Byzantines and Chinese to his court to decorate the walls in one of his chambers. This enabled him to compare which of them were better. The Byzantines asked him to provide all kinds of paint and tools so they could produce their masterpiece, while the Chinese only asked for materials needed for cleaning and polishing. Once the work was completed, the Byzantines' project was unveiled and enchanted everyone. Then they looked toward the artists from China, who were standing in front of a giant curtain concealing

their part of the room. When the curtain was pulled aside, the enchantment was even greater: A huge mirror reflected a perfect image of the Byzantines' artwork.

Thus, whereas theorists ("the Byzantines") progress toward excellence by adding many wonderful things, sciences and virtues to their minds, Sufis ("the Chinese") achieve excellence by cleansing their soul like a mirror, which is a typical theme in Muslim philosophy and spirituality. The content and knowledge they attain, however, is exactly the same. The conclusion from the story is that what decides about the difference in the lifestyle and knowledge a given person has is, above all, that person's inner attitude, their approach to their own mind. Of course Al-Ġazālī underlines the fact that the philosophers' undoing is their pride and their belief that science alone can be enough for their salvation and happiness. According to the author, meanwhile, the most important thing is not so much to consider the concepts of sciences, which are one of the tools for improving the soul by enabling it to gain access to knowledge, as to define the path to happiness which, contrary to falsafa, can only lead through practice, to the exclusion of sciences (*Mīzān al-ʿamal*, p. 45; cf. also: *Iḥyāʾ ʿulūm al-dīn*, vol. 2, pp. 1379–1382).

As Al-Ġazālī underlines, philosophers find it especially hard to listen to wise guides and to doubt their own beliefs. In a letter to his followers he writes directly that

> giving advice is simple. The problem lies in practicing it, because its taste is bitter (Arabic: *ḏawq murr*) to those who follow their desires, because their heart is close to misdeeds. This applies particularly to [people] studying theoretical sciences that deal with the urges of the soul (Arabic: *ḥaḏ al-nafs*) and worldly goods (Arabic: *manāqib al-dunīā*). They presume that pure knowledge, which requires no practice, is the way to salvation and rescue. That is the opinion of the philosophers. Glory be to God, the Almighty! Fools such as they do not know that by acquiring knowledge but not practicing it, they are preparing the harshest judgment for themselves. Because as the Prophet said, (peace and blessing be upon Him): 'on Judgment Day they will suffer the most whom God entrusted with knowledge, but they did not take advantage of it' (*Ayyuha ʾl-Walad*, pp. 91–92; in: Reisman, 2010).

Al-Ġazālī was aware of the existence of a philosophical way of life, but in his view it resulted rather in a life of scientists detached from reality or exploiting their position to enjoy

worldly pleasures. "Even if you pour two thousand bottles of wine but do not drink them, you will not be drunk" (*Ayyuha ʾl-Walad*, p. 92). In other words, even if you learn thousands of philosophical arguments but do not live by them, you will not achieve wisdom – this is the advice, Socratic even, we find in Al-Ġazālī's classic letter. On the other hand, it needs underlining that it was not necessarily doctrinal errors that were the main kind of error made by philosophers. This is especially noticeable in the treatise *Criterion of Action*, in which the work of "theorists" – the Byzantine artists – is not presented in dark colors. After all, the king invited not only the Chinese (Sufis) but also the Byzantines (philosophers) to his court. Al-Ġazālī also gave a metaphorical description of the beauty that comes from the work of the philosophers. The most serious accusation seems to be that the philosophers do not necessarily treat seriously the relationship between theory and study on one hand and practice and life on the other. This is exactly why Al-Ġazālī saw a necessity to develop his own, new concept of science, which he did in the treatise *The Revival of the Religious Sciences*.

In his magnum opus Al-Ġazālī outlined three fundamental branches of knowledge: the "science of praxis" (Arabic: *ʿilm al-muʿāmala*), the "science of unveiling" (Arabic: *ilm al-mukāšafa*) and, the most important of them, the "science of the hereafter" (literally: "science of the path to the hereafter"; Arabic: *ilm ṭarīq al-āḫira*). Insofar as the "science of praxis" and the "science of unveiling" echo the philosophical division into "practical" and "theoretical" sciences (cf., among others, Gilʿadi, 1989; Treiger, 2012), the "science of the hereafter" is a new term. The task of this science is to designate a goal, namely ultimate happiness, which the previous two sciences, in Al-Ġazālī's view, are unable to designate (being devoted exclusively to either theory or practice): "the science (Arabic: *ʿilm*) thanks to which we strive for the goal of the hereafter (Arabic: *al-ʿilm allaẓī yutawaġġahu bihī ilā al-āḫira*) is divided into the 'science of praxis' and the 'science of unveiling'. The science of unveiling is aimed at revealing the object of cognition (Arabic: *kašf al-maʿlūm faqat*), whereas the science of practice, besides revealing it, is aimed at acting according to it (Arabic: *al-ʿamal bihī*)" (*Iḥyāʾ*, vol. 1, p. 3).

What causes us to practice the "science of the hereafter" well? What distinguishes proper from improper practice? According

to Al-Ġazālī, the fundamental difference lies precisely in an emphasis on spiritual practice connected with exercising the consciousness and rejecting "pure" science (i.e. based solely on theory or reasoning) as an effective tool for attaining happiness. On the other hand, in this way Al-Ġazālī carries out a "rationalization of Sufism" (Abrahamov, 2015, p. 37), using Sufi terms to express a concept that is actually close to the philosophers' way of thinking. In fact it is a reasoned, rational basis that is able to prevent "the poisoning of rational power. [Without it], someone could say 'I am the truth!' and another 'glory be to me, I am so great!' … That is why the power of reason is God's measure in His world" (*Miškāt al-anwār*, p. 18). Al-Ġazālī is aware that one should not exercise the soul without an appropriate "measure" – a tool ensuring some degree of objectivity of judgment and protecting one against madness, which seems to have often overwhelmed Sufi masters[4]. The author of *The Revival of the Religious Sciences* thus moves away from a "mystic" understanding of imitation of God as deification, i.e. becoming wholly like Him in nature, toward a more philosophical concept of "imitating God's actions" by studying them.

5. Rejection of philosophy?

Does Al-Ġazālī reject all of philosophy, then? It is worth noting in this context that it is not without reason that Awerroes calls Al-Ġazālī a "chameleon" who "is an Ashʿarite to Ashʿarites, a Sufi to Sufis, and a philosopher to philosophers" (after: Hourani, 1985, pp. 135–136). As Treiger has shown, regarding many issues Al-Ġazālī drew on the philosophy of Avicenna, directly invoking his theory of the soul, intellect, his concept of ethics and theology – only calling them by different terms (Treiger, 2012, p. 6 and 119, p. 34). Al-Ġazālī states outright that acquiring knowledge – also using philosophical tools – is a way not only of getting to know God (understood as the fullness of truth), but also of imitating Him (Arabic: *tašabbuh*) and becoming more like Him (Arabic: *taʿalluh*; cf. *al-Maqṣad*

---

[4] Al-Ġazālī is referring directly to a Sufi master, Al-Ḥallāǧ (d. 310/922), who is famous for his saying "I am the truth," for which he was sentenced to death by hanging. This pronouncement was meant to express his full union with God, the path leading to whom was excellence of practice.

al-asnā, p. 65). Philosophical study is also a good way of "taming" the irrationalism that is associated with excessively performed religious practices.

Importantly, however, according to Al-Ġazālī not all concepts of the philosophers were generally contrary to the rules of religious law. One example applying to people was "the view that one can rationally justify that the soul is a substance existing in and of itself (Arabic: *al-nafs ǧawharān qāʾimān bi-nafsihi*)" (*Tahāfut al-falāsifa*, p. 181). As the author of *The Precipitance* goes on to explain, this means above all "their view that reason itself is able to justify this and that there is no need to resort to revelation" (*Tahāfut al-falāsifa*, p. 182). On the other hand, in his opinion there has to be a foundation external to the human soul enabling the construction of a knowledge structure and, consequently, life practice leading to happiness. That foundation is God, whose existence prevents "absurd" conclusions regarding the possibility of existence of "knowledge about knowledge about knowledge etc." (cf. Griffel, 2009b). According to the author of *The Precipitance of the Philosophers*, human knowledge needs an external foundation, otherwise it will be necessary to accept that it is its own foundation, and that leads to regression ad infinitum.

It also needs mentioning that in a few of the discussions in *The Precipitance*, Al-Ġazālī mentions "refutation" (Arabic: *ībṭāl*) of various arguments, but – as Griffel points out – refutation does not mean rejecting the philosophical doctrines analyzed in this book: "It is clear that in his *Incoherence* Al-Ġazālī does not set out to prove the falsehood of all of – or even of most of – the philosophical teachings discussed there" (Griffel, 2009b, p. 98). Meanwhile, Treiger points out that the term *ībṭāl* in this case means criticism of the reasoning and not its conclusions, most of which Al-Ġazālī accepts (Treiger, 2012, p. 94):

> It is in this sense that Al-Ghazālī's *Precipitance* can be described as a 'pseudorefutation,' a kind of exercise in deconstructionist rhetoric and dialectic … *The Precipitance* operates on the level of *iʿtiqād* (opinion) rather than *maʿrifa* (cognition) and is written for the *mutakallimūn* [representatives of *kalām*], about whom Al-Ghazālī always speaks with disdain, numbering them among the 'common folk' (*ʿawāmm*). *The Precipitance* is thus essentially a work of *kalām*, aiming at safeguarding the commoners' creed (*ʿaqīdat al-ʿawāmm*), nothing more and nothing less (Treiger, 2012, p. 95).

Jules Janssens was one of the authors quickest to notice the problematic nature of standard interpretations that saw the "refutation" or rejection of philosophy, or at least uncompromising polemics with falsafa, as the primary and real purpose of Al-Ġazālī's work. According to this scholar, "[*Tahāfut*] should most probably be classed among works of the genre of *taliqa*, i.e. an advanced student's work which presents a straightforward commentary … of a writing of the master" (Janssens, 2001, p. 1).

To conclude, it is worth mentioning that the most important, strictly religious work by Al-Ġazālī, *The Revival of the Religious Sciences*, was publicly condemned and burned in Cordoba in 1109 (two years before his death), and the accusations presented for the first time in Nishapur were repeated on this occasion. In later centuries many Muslim law experts and scholars warned against reading Al-Ġazālī, as an author "poisoned" by Avicenna and Sufism. Today, however, staying in Muslim countries it is easy to see it is actually Al-Ġazālī's works that occupy prime places in libraries and bookstores, placed on the shelves right next to the Qur'an. Al-Ġazālī has been given the title "Ḥuǧǧa al-īslām" – "Proof of Islam" – and regardless of various controversies, today remains one of the most important Muslim writers in history. This fact is especially interesting in the context of his ambiguous attitude toward philosophy and the way he combined it, including spiritual practices known to the philosophers, with an orthodox way of life in Islam.

References

Abrahamov, B. (2015). Al-Ghazālī and the rationalization of Sufism. In: G. Tamer (Ed.), *Islam and Rationality. The Impact of al-Ghazālī. Papers Collected on His 900th Anniversary* (pp. 35–48). Leiden: Brill.

Abul Quasem, M. (1974). Al-Ghazali's rejection of philosophic ethics. *Islamic Studies*, vol. 13, 2, 111–127.

Al-Ġazālī (1356/1937–57/38). Iḥyāʾ ʿulūm al-dīn [Revival of the Religious Sciences]. In: *Iḥyāʾ ʿulūm al-dīn*, Laǧnat Našr al-Taqāfa al-Islāmiya, vol. 1–4, Cairo. Retrieved from www.ghazali.org/site/ihya.htm

Bīǧū, M (Ed.) (2008). *Mīzān al-ʿamal*. Damascus: Dār al-Taqwā.

Buchman, D. (Ed.) (1998). *al-Ghazālī. The Niche of Lights*. Provo: Brigham Young University Press.

Garden, K. (2014). *The First Islamic Reviver. Abū Hāmid al-Ghazālī and His Revival of the Religious Sciences*. Oxford: Oxford University Press.

Gilʿadi, A. (1989). On the origin of two key-terms in Al-Ġazzālī's *Iḥyāʾ ʿUlūm al-Dīn*. *Arabica,* vol. 36 (1), 81–92.

Griffel, F. (2009). *Al-Ghazali's Philosophical Theology*. Oxford: Oxford University Press.

Griffel, F. (2009b). Al-Ghazālī's cosmology in the veil section of his *Mishkāt al-Anwār*. In: T. Langermann (Ed.), *Avicenna and His Legacy: A Golden Age of Science and Philosophy* (pp. 27–49). Turnhout: Brepols.

Günther, S. (2010). The principles of instruction are the grounds of our knowledge: Al-Faārābi's (d. 950) philosophical and al-Ghazāli's (d. 1111) spiritual approaches to learning. In: O. Abi-Mershed (Ed.), *Trajectories of Education in the Arab World: Legacies and Challenges* (pp. 15–35). London: Routledge.

Hourani, G. (1985). *Reason and Tradition in Islamic Ethics*. New York: Cambridge University Press.

Jabre, F. (Ed.) (1969). *Erreur et délivrance*. Beirut: Commission libanaise pour la traduction des chefs d'œuvre.

Janssens, J. (2001). Al-Ghazzālī's *Tahāfut*: Is it really a rejection of Ibn Sīnā's philosophy? *Journal of Islamic Studies*, vol. 12 (1), 1–17.

Kukkonen, T. (2016). Al-Ghazali on error. In: F. Griffel (Ed.), *Islam and Rationality. The Impact of al-Ghazālī. Papers Collected on His 900th Anniversary*, vol. 2 (pp. 3–31). Leiden: Brill.

Linant de Bellefonds, Y. (2012). Fāsid wa Bāṭil. In: P. Bearman, Th. Bianquis, C.E. Bosworth, E. van Donzel, & W.P. Heinrichs (Eds.), *Encyclopaedia of Islam. Second Edition*; URL: http://dx.doi.org/10.1163/1573-3912_islam_COM_0215

Marmura, M. (Ed.) (2000). *The Incoherence of the Philosophers*. Provo: Brigham Young University Press.

Menn, S. (2003). The discourse on the method and the tradition of intellectual autobiography. In: J. Miller, B. Inwood, *Hellenistic and Early Modern Philosophy* (pp. 141–191). New York: Cambridge University Press.

Reisman, D. (Ed.) (2010). Al-Ġazālī. Ayyuhaʾl-Walad [O, Son!]. In: *Classical Foundations of Islamic Educational Thought: A Compendium of Parallel English-Arabic Texts* (pp. 90–107). Provo: Brigham Young University Press.

Rosenthal, F. (2014). *Man versus Society in Medieval Islam*, D. Gutas (Ed.). Leiden: Brill.

Shehadi, F. (1971). Al-Ġazālī. Al-Maqṣad al-asnā fī šarḥ maʿānī asmā Allāh al-ḥusnā [The Highest Purpose: Explaining the 99 Names of God]. Beirut: Dar el-Mašreq.

Treiger, A. (2012). *Inspired Knowledge in Islamic Thought. Al-Ghazali's Theory of Mystical Cognition and Its Avicennian Foundation*. London: Routledge.

Maria Łukaszewicz-Chantry
ORCID 0000-0002-1920-6848
University of Wrocław

Only a Poet Never Lies…
Maciej Kazimierz Sarbiewski's Thoughts on the Privilege of Poets

Abstract

In his treatise *De perfecta poesi, sive Vergilius et Homerus*, Maciej Kazimierz Sarbiewski presents his reflections on poetic art and states that poets never lie. This privilege results from the conviction that a poet creates in the same way as God, bringing his characters to life. That is why "he can speak about what is not, as if it really existed." Poetic fictions are also often a veil behind which a precious truth is hidden. This truth can be reached by using an appropriate allegorical interpretation.

Keywords: Mimesis, probability, poetic art, fiction, allegoresis

In *De perfecta poesi*, *sive Vergilius et Homerus*, a treatise dedicated to the epics, Maciej Kazimierz Sarbiewski SJ (1595–1640) wrote:

> Only a poet never lies, because even if he says something that does not agree with the truth, he does not say it with the intention of being believed, but so that his work be read as the perfect expression of a certain idea, and that the conclusion drawn from the discovered details show the deeper truth hidden beneath his tale (Sarbiewski, 1954, p. 14).

Where does this privilege come from? How is one to understand it, considering that it is claimed by an author who was himself a well-known and valued poet in 17th-century Europe, where he was called by the prestigious name of the Christian or Sarmatian Horace?

1. *Poeta fingit*

To answer this question, let's take a look at the original Latin sentence:

> Solus poeta numquam mentitur, nam etsi ea dicat, quae non sunt, non dicit tamen eo animo, ut credantur, sed ut cognoscantur veluti perfecte expressa, ut excognitis ulterior quaedam veritas sub fabula latens colligatur (Ibid.).

The first part of the sentence says, literally: "Only a poet never lies, **even if he says things that are not, … .**"

The understanding of a lie suggested by Sarbiewski may then be associated with Plato's reflection on truth and falsehood. According to the Greek philosopher, a false statement is one that "says about things which are not, as if they were" (Plato, *Sophist* 263a, translated by B. Jowett). Aristotle defined truth and falsehood similarly in his *Metaphysics*:

> To say of what is that it is not, or of what is not that it is, is false, while to say of what is that it is, and of what is not that it is not, is true (Aristotle, *Metaphysics* IV, 7, 1011b 25–28, translated by W. D. Ross).

Why, then, according to Sarbiewski, can a poet present things that are not, as if they were? Is the intention itself sufficient, i.e. that he does not want to mislead anyone and that he does not expect to be believed?

Here, we come to the issue of poetic creation and the eternal dispute between philosophy and poetry. Plato accused poets of being confusing in their works (*Republic* 605 D-E). He even wanted to exclude them from his Ideal State. He believed that poets create images which imitate empirical reality, while reality itself is a reflection of an idea. The mimetic arts, poetry among them, were thus "copies of copies," moving away from the world of perfect ideas, and therefore from the Truth, which is the most important of them. Plato accused poets of creating a vision which is likely to be treated as real when the recipient lets himself be misled by the illusion. What's more, this vision affects the irrational part of the soul, causing emotional confusion.

Plato's views on poetry were refuted by Aristotle, who claimed that poetry does not directly imitate the empirical reality, but, based on the principle of probability, creates what is possible or

necessary. A poet can therefore present something that does not exist, but could or should exist: He does not relate "… what has happened, but what may happen, — what is possible according to the law of probability or necessity" (Aristotle, *Poetics* IX 1451a, 35, translated by S. H. Butcher).

A poet creates a kind of "parallel world" (*analogon*) which is not a "copy of a copy" anymore, but an independent universe. Therefore the criterion of truth as conformity with reality becomes inadequate in the case of poetry. Significantly, Aristotle even accepts the introduction of wonderful elements (*mirabilia*) to certain genres, as long as they are consistent with the logic of the plot (Ibid., XXIV 1460a, 10–15). He also emphasizes that poetry tends to express the universal, while history tells the particular (Ibid., IX). With Aristotle, poetry becomes closer to philosophy.

In the Renaissance and Baroque, when writing about *mimesis*, the authors of poetic treatises relied on the authority of Aristotle, although they sometimes tried to find new interpretations of his theory. Sarbiewski also follows the Stagirite in his own reflections on poets, creative power and poetry – an art which occupies a special place among the mimetic arts. According to the Jesuit, "poetry is an art that imitates beings by using verbal material, not according to how they exist, but how they should or could exist or have existed" (Sarbiewski, 1954, p. 4).

The term "poet" itself reflects well the nature of literary creation. Sarbiewski writes that the ancient Greeks were right to seek the origin of this word in the verb *poiein*, which is of "undecided meaning" (*dubia significatio*) as it expresses both production (*faciendum*) and reproduction (*rursus fingendum*), which means imitation (*imitandum*) (Ibid., p. 3). Other authors of late Renaissance poetic treatises also dealt with the semantic analysis of the verb *poiein*, e.g. Jacobus Pontanus (Cf. J. Pontanus, *Prawidła poetyckie*, in: *Poetyka okresu renesansu. Antologia*, 1982, p. 485; cf. Janus, 2006, pp. 63–64), who may have been a source of inspiration for the Polish Jesuit.

Sarbiewski himself, when writing about the work of a poet, most often uses the verb *fingere* (form, shape, create, invent, imagine, pretend), not only in his *De perfecta poesi* treatise but also in other writings. For example, in *Characteres lyrici* (a kind of lyrics textbook), he writes that Horace imagines (*fingit*) himself hovering like a bird over the earth (Sarbiewski, 1958, pp. 27–28; cf. Łukaszewicz-Chantry, 2002, pp. 11–50). Such

peculiar "fictions" (*fictiones*) are typical for lyrics, while in the epics, "fictions" should be more common and probable (Sarbiewski, 1958, pp. 29–30). Lyrical fictions may even be "contrary to nature" (*contra naturam*), and the more unusual and "imaginary" (*abstractae*) they are, the more they testify to the talent of the poet. Fictions, therefore, do not mislead the reader: They are a natural expression of creative invention, a certain aesthetic vision intended to be pleasant to the audience. For "who, for example, ever saw a flying poet?" (Ibid., p. 30).

The verb *fingere* is also frequently used in Renaissance poetic treatises in connection with poets' fictional creativity. Authors of humanistic poetics usually combine *fingere* and *imitari* (invention and imitation), and they do not see contradictions between them, but a dynamic complementarity (Niebelska-Rajca, 2011, pp. 103–105). So does Sarbiewski when he writes that a poet simultaneously creates and imitates. However, he puts special emphasis on creativity as a distinctive feature of poetic art.

2. *Instar Dei poeta*

In order to explain that creation and imitation combine with each other and complement each other, the author of *De perfecta poesi* uses a major argument: This is exactly how God himself creates. He creates man in his image and likeness, which means that in the act of creation, he imitates himself. Sarbiewski finds even more analogies between a poet's and God's way of creating (Cf. Sarnowska-Temeriusz, 1967, pp. 129–130; Stawecka, 1989, pp. 74–94). While other artists already have existing material and use it to shape their works (e.g. a sculptor), a poet invents his characters and their actions himself from the beginning. He brings to life both the topic and the way it is shaped (*et materiam condere, et formam rerum*) (Sarbiewski, 1954, p. 2).

A poet, like God, creates by using words. Similarly, he brings his characters to life by naming them (*poetae est dare nomina*), like God, who in the act of creation gives names to his creatures. Referring to the authority of Saint Paul, Sarbiewski writes:

> … in this, likewise, a poet is, in a sense, similar to God (*similis Deo poeta*) who, when He creates something – according to Saint Paul – *calls into existence things that do not exist,* and so gives shape to what did not exist before (Ibid., p. 4).

Therefore when a poet presents "what is not, as if it were," he does not mislead his audience; what's more, this is actually what poetic creation is about: "giving some new existence" (Ibid., p. 2). Sarbiewski emphasizes that historians and orators are likely to lie, while a poet "has the privilege to speak about things that are not, as if they really existed, and even to pretend that they exist, and doing so, not to lie at all" (Ibid., p. 4).

Giving further consideration to a poet's special affinity with the Creator, Sarbiewski also explains that it is God who is the driving force of all poetry, including pagan poetry (Ibid., p. 11). Belief in the divine sources of poetry dates back to ancient times. Plato himself wrote about poetic inspiration (*furor poeticus*) sent by the gods. His ideas inspired Renaissance humanists, especially Neoplatonists. For Sarbiewski as a Christian, poetic inspiration is sent by the Holy Spirit, and this allows him to draw further conclusions: Since a poet is an instrument of the Holy Spirit, he may – although sometimes not quite consciously – pass on some truth in the form of fiction. This idea is formulated by the author of *De perfecta poesi* in the second part of the sentence quoted at the beginning of this paper:

> … and that the conclusion drawn from the discovered details show the deeper truth hidden beneath his tale *(ut excognitis ulterior quaedam veritas sub fabula latens colligatur)* (Ibid., p. 14).

In his invitation to look for a second, deeper layer hidden in poetry, Sarbiewski invokes the hermeneutic method of the ancient Greeks. Theagenes of Rhegium (6th century BC) was the first to suggest that the Homeric gods' petty quarrels and fights should be interpreted in the spirit of Ionic philosophy of nature as a battle of the elements. With time, allegoresis became a compromise solution to the conflict between philosophy and poetry.

Other Christian authors preceding Sarbiewski also recommended a search for hidden truth in mythology and poetry, to mention Albertino Mussato from Padua (1261–1329) who said that poetry, also pagan verse, is "a science sent from heaven." Greek myths convey revealed truth, but in a hidden way (*enigme*) (Cf. Łukaszewicz-Chantry, 2019, pp. 181–182). Petrarch and Boccaccio expressed similar beliefs: To them, the first poets were theologians who transmitted truth in the form of beautiful fiction. Sarbiewski follows them when he claims that mythology is,

moreover, a *theologia fabulosa* (from *fabula*, fable, fiction), because the pagans had only a "dreamy knowledge" of the One God, which they expressed through their images and myths (Sarbiewski, 1972, passim; cf. Urbański, 2000, pp. 15–37; Łukaszewicz-Chantry, 2019, p. 182). He also believes that for some poets, the truth could be kept hidden intentionally, like a priceless treasure (Sarbiewski, 1972, p. 223). However, it is possible to discover this hidden truth using the right key, which is the application of allegoresis (Cf. Sarnowska-Temeriusz, 1967, pp. 144–146).

Sarbiewski writes that both historians and orators can convey truth. A poet, however, does it in a nobler way (*nobiliorem modum*), because "he does not present the naked truth (*nudam veritatem*), ... but shows it clad in beautiful garments (*vestitam honeste*), as if covered with a cloth of fable (*quasi pallio fabulae tectam*), like the merchants who used to sell their valuable items by presenting them under a veil" (Sarbiewski, 1954, p. 16).

Therefore the poet not only does not lie when he "fantasizes" (*fingit*) to create a fable: His poetic "fictions" (*fictiones*) can be a way of passing on a truth, especially a precious truth which would be difficult to express directly.

References

Aristotle (2008). *Poetics*, transl. S. H. Butcher. *The Project Gutenberg EBook of Poetics, by Aristotle.* Retrieved from https://www.amherst.edu/system/files/media/1812/The%252520Poetics%252520of%252520Aristotle%25252C%252520by%252520Aristotle.pdf.

Janus, K. (2006). Wokół pojęcia twórczości. Ze studiów nad *De perfecta poesi* Macieja Kazimierza Sarbiewskiego [Around the notion of creative output. From studies on Maciej Kazimierz Sarbiewski's *De perfecta poesi*]. *Prace naukowe Akademii im. Jana Długosza w Częstochowie*, vol. X, 59–67.

Łukaszewicz-Chantry, M. (2002). *Trzy nieba. Przestrzeń sakralna w liryce Macieja Kazimierza Sarbiewskiego* [Three Heavens. Religious Space in the Poetry of Maciej Kazimierz Sarbiewski]. Wrocław: Wydawnictwo Uniwersytetu Wrocławskiego.

Łukaszewicz-Chantry, M. (2019). *Ciceronianus* czy *Christianus*. Dawni autorzy chrześcijańscy wobec literatury pogańskiej [*Ciceronianus* or *Christianus*. Early Christian authors' approach to pagan literature]. *Więź*, 1, 176–182.

Niebelska-Rajca, B. (2011). *Poeta imitatore* czy *poeta facitore*? Późnorenesansowe włoskie dyskusje o *mimesis* [*Poeta imitatore* or *poeta facitore*? Late Renaissance Italian discussions on mimesis]. *Odrodzenie i Reformacja w Polsce*, LV, 101–122.

Plato (2013). *Sophist*, transl. B. Jowett, *The Project Gutenberg EBook of Sophist, by Plato*. Retrieved from http://www.gutenberg.org/files/1735/1735-h/1735-h. htm.

Ross, W.D. (Ed. & transl.) (1912). Aristotle's *Metaphysics*. Oxford: Clarendon Press.

Sarbiewski, M.K. (1954). *O poezji doskonałej, czyli Wergiliusz i Homer* (Polish edition of *De perfecta poesi, sive Vergilius et Homerus*, S. Skimina (Ed.), transl. M. Plezia). Wrocław: Ossolineum.

Sarbiewski, M.K. (1958). *Wykłady poetyki* (Polish edition of *Praecepta poetica*, transl. S. Skimina). Wrocław: Ossolineum.

Sarbiewski, M.K. (1972). *Bogowie pogan* (Polish edition of *Dii gentium*, transl. K. Stawecka). Wrocław: Ossolineum.

Sarnowska-Temeriusz, E. (1967). Teoria poezji Macieja Kazimierza Sarbiewskiego [Maciej Kazimierz Sarbiewski's theory of poetry]. In: M. Głowiński (Ed.), *Studia z teorii i historii poezji* [Studies from the Theory and History of Poetry] (pp. 126–147). Wrocław: Ossolineum.

Sarnowska-Temeriusz, E. (Ed.) (1982). *Poetyka okresu renesansu. Antologia* [Poetics of the Renaissance: An Anthology]. Wrocław: Ossolineum.

Stawecka, K. (1989). *Maciej Kazimierz Sarbiewski prozaik i poeta* [Maciej Kazimierz Sarbiewski: Prose Writer and Poet]. Lublin: Wydawnictwo Towarzystwa Naukowego KUL.

Urbański, P. (2000). *Theologia fabulosa. Commentationes Sarbievianae*. Szczecin: Wydawnictwo Naukowe Uniwersytetu Szczecińskiego.

Izabella Zatorska
ORCID 0000-0003-2048-4033
University of Warsaw

Illusion and Truth in Theater from the Baroque to Romanticism

Abstract

What does theater, as a metonymy for literature and art, represent in the centuries of transition? French classical theater was aware of difficulties in using the Aristotelian rule of mimesis. The problem of the impossibility of capturing the essence of things, when post-truths were not yet known, was acknowledged, in theory and in practice, by Baroque theater: From the perspective of *theatrum mundi*, with the use of "theater within the theater" (Shakespeare, Pierre Corneille, Molière), it showed the hidden, disturbing depth of "the norm of the day." The problem of appearances (Lesage, Marivaux) resulted in the aesthetic and moral reform of Diderot, theoretician and practitioner of *drama*. Radical in their revolt, the protagonists of Schiller and Musset rejected both any compromise with illusion and the possibility of reforms. At the same time, they longed for the incarnation of the ideal of pure love, and were prepared to destroy others and even themselves if this ideal was to turn out to be a lie. Goodness and responsibility, as the basis for a play, are the values which – according to Father Józef Tischner – promise that Abel's choice will prevail over Cain's.

Keywords: truth, illusion, theater within the theater, mise en abyme, mimesis

The uniqueness of theater as a medium that is dual by definition – stage vs. audience, performance vs. reality – prevents it from offering any direct treatment of axiology in relation to the truth[1]. In an era admitting the possibility that (solely) post-truths exist, we need to go back to writers distant in time in order to find the will to seek and define truth and, ultimately, to defend – even at the cost of lives – all that has been

---

[1] Georges Forestier points out that the communion between actors and audience, sanctified by the content of performances (on religious themes) was severed in the early 16th century (Forestier, 1996 [1981], p. XI).

acknowledged as truth. This applies to stage characters placed within binomial structures (truth-falsehood, right-wrong, seriousness-laughter) which – building distance toward the issues being presented – give audiences and readers of a play the privilege of free judgment. Baroque writers – Shakespeare, "young" Pierre Corneille, Molière sometimes – played with illusion using the "theater within the theater" concept. It is sometimes hard to distinguish it from the play of mirrors, or the self-referential *mise en abyme* that Marivaux applied in the 18th century, using the aforementioned doubling of the stage and audience less literally; we could count it among the Rococo period's "double register" experiments[2], more serious in their intention than is widely believed about this style. Rejuvenating the aspirations of neoclassical mimesis that Lesage took over from Molière (their two "Messieurs T" – Tartuffe and Turcaret – simply beg comparison), Diderot placed not tragedy or comedy at the focus of theatrical ambitions, but *drama*. Despite his ambition to get as close as possible to the audience, the "domestic tragedies" and "serious comedies" of the author of *Jacques the Fatalist* were only moderately successful. Theater of the French Revolution harnessed truth in the service of ephemeral ideologies of consecutively collapsing political groups – but this is not the kind of theatricalization of the issue of truth and illusion we are considering here. The romantic (sometimes slightly melodramatic) characters created by Lessing and Schiller (great admirers of Diderot's ideas) saw the choice between the world's lies and their personal ideals as a make-or-break decision. But what if the truth, as encyclopedist Diderot suggested, consists in traveling toward it – in the infinity of a finite biography? The bitterness involved, especially when you resort to fraud, like in the "armchair" productions of Alfred de Musset[3], again makes the prospect of capturing the essence of the matter more distant. What conclusions from the building of relationships between truth and illusion can be drawn based on a mini-fragment from the history of European theaters?

---

[2] A term from Jean Rousset's study *Forme et signification. Essais sur les structures littéraires de Corneille à Claude*, 1962.

[3] The title he gave a collection of his plays from 1832 was *Armchair Theater – Un Spectacle dans un fauteuil*.

1. Arguing about mimesis:
torture as therapy?

In the introduction to the monographic volume *Mimesis w literaturze, kulturze i sztuce* [Mimesis in Literature, Culture and Art], Zofia Mitosek presents different applications of the principle from the title (magical, expressive, sociological, psychological mimesis), often as part of a game – whether social or psychological; Greek theoretical thought in the persons of Plato and Aristotle was the first to ask about truth, i.e. "the relationship between the imitating object and the reality being imitated." The relational criterion remained relevant in modern French literary theater – produced on the stage of the Comédie Française – right up to the Revolution. However, insofar as "mimetic art in Plato's concept was a lie," Aristotle "recognized that an imitative work of art reflects not actual but probable reality. … In such a situation, mimesis is the process of imitating that which can be or should be, and not that which is" (Mitosek, 1992, pp. 22–23). Especially the kind of imitation that causes audiences to feel two passions – pity and fear – toward an innocent character, hurt by fate and bearing similarities to the spectator, serves the purpose of psycho-moral purgation (*kàtharsis*) at the level of intellect and feelings (Podbielski, 1989, p. LXXXIV) – expanding the notion of *kàtharsis* itself to include new meaning (besides medical, ritual and intellectual – cleansing from toxins, sins and errors; ibid., pp. LIX-LXIV).

Building a distance toward represented reality is certainly facilitated by a form that initially was used as an interlude but, in modern-age England and France, gradually "colonized" plays: theater within the theater. At first Georges Forestier defined it as a structure that grew from the aesthetics of the Baroque, which was conducive to games of illusion; in 1996 he acknowledged that it also appeared in other periods. To put it as simply as possible, theater within the theater (twt) involves "stopping the dramatic plot by incorporating an autonomous element that the characters perceive as a performance (*du théâtre*)" (Forestier, 1981/1996, p. XI). The beginnings of twt appeared in antiquity (the Greek chorus) and almost disappeared in the Middle Ages (the prologue was its relic). The first modern examples of this structure surfaced in Portugal (1532), in England (1589), and in *commedia dell'arte* which served as a mediator helping spread

the practice in different countries. Insofar as, across the Channel, Shakespeare was not averse to twt, in France – if we agree with Forestier from 1981 that the height of twt fell on the years 1628–1694 – it appeared, at best, in second-rate plays by the greatest masters. Among these we can count the intermedia in Molière's *The Bourgeois Gentleman* (1670) and in *The Imaginary Invalid* (1673) (ibid., pp. 10–11).

Jolanta Dygul distinguishes "three types of theater-within-the-theater structures in 17th-century Italian drama: 1) *mise en abyme* (self-referential), i.e. introducing "an inner play into the play by means of analogy or parody"; 2) a pretext play whose "conventionalized character reinforces the illusion of 'reality' in the framework play" because it "reinforces the difference between the plot levels, attracting the audience's attention with familiar and well-liked forms. It has no connection to the plot of the main play"; 3) "a theater rehearsal: The framework play portrays the life of an acting company or the organization of a theater, and the inner play – the performance being prepared" (Dygul, 2010, p. 66). Shakespeare's *A Midsummer Night's Dream* (1595 or 1596, published 1600), in which a troupe of Athenian craftsmen rehearses a tragedy, or rather "the most lamentable comedy and most cruel death of Pyramus and Thisbe," fulfills types 1 and 3 simultaneously. The play performed by Peter Quince and company for the wedding festivities of Theseus, prince of Athens, and Hippolyta, former Amazonian queen, thematizes the difficulties experienced by a couple from the Athenian nobility, Hermia and Lysander, suggesting their "alternative, less than happy end, thus underlining the conventional character (i.e. accepted by both the author and the audience) of this genre of drama and art in general" (Jankowski, 1992, p. 8). In a paradox, the viewers (both inside – characters in the play, and outside – the audience) accept the possibility of sober judgment of reality coexisting with infatuation. If we consider the play through the anthropological analysis model that Katarzyna Wielechowska borrowed from René Girard (Wielechowska, 2013, pp. 105–125), we can go beyond the analogies, play of mirrors, multiplication of registers (pathos, irony, parody, grotesque) mentioned by Polish Shakespearologist Jan Kott (Wielechowska, 2013, pp. 107–108). Distorted into mimetic desire symbolized by a "love triangle" (repeated in different combinations, in the worlds of the craftsmen, the nobility and the elf spirits), mimesis still shows the

truth – not metaphysical, not prehistoric, but pre-mythical truth: "Shakespeare proposes a mimetic theory of mythical origins," as Wielechowska quotes Girard, while the elf-spirit subplot "is needed to show the process of mythological distortion" (ibid., p. 114). Parallel to the anthropological, social plane, twt makes possible a metatheatrical and metalinguistic approach to the relationships of truth-illusion-falsehood and fiction-reality.

Shakespeare's culturally syncretic space (Midsummer Night in ancient Athens precedes the wedding of Theseus and Hippolyta, queen of the Amazons) can be compared to the fairy-tale-like and phantasmagorical space of Marivaux's *The Argument*, a one-act play from 1744. A century and a half after Shakespeare's metatheatrical play, Marivaux created a version built into the framework of a philosophical "dispute," i.e. a debate which, in the world of his comedy, had taken place twenty years before, at a court, regarding who committed the sin of infidelity first: man or woman. The experiment ended in a draw: Of the four young people raised in solitude in a forest location (*enceinte*), each would, to some extent, give in to… is it not actually that same "mimetic desire" traced by Girard? As Wielechowska notes after him, it embroils "the characters in a mimetic process eliminating the differences between them and transforming them mutually into each other's doubles" (ibid., p. 112). In Marivaux's play the cause of this process is diagnosed quite accurately: Access to pre-mythical truth about the beginnings is obscured by contemporary awareness and custom. The assumed "natural" education of the four children is meant to avoid osmosis with courtly culture, hence the choice of two black servants as the teachers. However, their behavior is a carrier of "worldly" culture: starting from language, stretching between a mirroring, narcissistic self and gender identity oriented toward a primal rivalry with women over men, all the way to the educational material (Carise, one of the servants, reminds Eglé about her music lessons) and interpersonal relations (the educators force the youngsters to take "brief absences" to prevent boredom with their partner). But in the perpetuum mobile driven by the four-fold ego's arrogance reduced to the sexual sphere, generating infidelity and obscuring the personal differences between the play's personae, Marivaux allows the characters to find a gateway toward freedom – not the leading foursome from the inner play (since the experiment unfolding in the woodland laboratory becomes theater within

the theater for the court), but a third couple, a kind of "control group." No mention of them was made in the Prince's prologue, i.e. the first two scenes of *The Argument*. But are Dina and Meslis precisely, who were not featured in the original script, not the voice of nature that the Prince mentions in scene two? *Deus ex machina*, in the form of *mise en abyme* – a thematized repetition of the princely couple's story in its optimistic version, saves the couple's chances for a future happy relationship; hence the Prince and Hermiane assure the inseparable lovers of their guardianship in future, whereas the ruler has the other four "put away" separately. In Marivaux's allegory, truth – is it metaphysical? – has psycho-moral rather than cognitive implications.

A very different role is played by twt in Pierre Corneille's *L'Illusion comique*, counted among the group of the Rouen-based playwright's first, irregular plays, next to *Place-Royale* or the famous *Le Cid*; we should call it a tragicomedy or a Spanish comedy, full of improbable adventures and mishaps, its complex composition putting it in close competition with the romance genre. Its first version was written in 1636, and a second one – more moralizing in the final intermedium – followed in 1660. Since just one "inner spectator" is enough to make us speak of an inner play, in Corneille's work that spectator is Pridamant, father of the spendthrift Clindor who – having grown weary of his father's tight rein – first found employment as an aide to a Gascogne captain, Matamore, and then, after abducting Géronte's daughter Isabelle, together with the girl he loves, her servant and the jailer's son (!), joined an itinerant group of actors, ultimately becoming the leader of a troupe that is a great success in the capital. The desperate father wants to find his son, and thanks to the magic of Alcandre is able to see past and – as he believes – current events of his only son's life. The latter events, which end in the death of someone who looks like Clindor, turn out to be nothing more than a play performed by actors. In Corneille's play the goal of arriving at the truth is achieved through a paradoxical method: The performance of a fictitious story – although perhaps not all is fiction, as it speaks of real threats to Clindor and Isabelle's relationship – is aimed at transforming Clindor's father in reality. In the conclusion, this hybrid comedy (marked by real death – that of Clindor's rival, as well as pretend death – that of Clindor himself as Theagenes, dissuading Rosine who loves him from being unfaithful; in the second

version of *L'Illusion* she was allegedly killed by her jealous husband) whose story "lasts as long as the show itself" – according to the playwright's foreword – offers praise of theater and the mission of an actor:

> … Now the theater / Is at a point so high that everybody idolizes it; / At one time treated with contempt, / Today it enchants all educated people, / It is the talk of Paris, dreamed about in the provinces. / It enjoys the patronage and affection of princes, / It is entertainment for the mighty, a delight to commoners, / It gives respite, fun, illusion; / And those whose deep wisdom is admired, / For whom the world and its people are their concern, / Find enough time to attend performances where / They forget at least for a moment about the burden of government. / Furthermore, our great king, whose majesty / Is feared at both ends of the world, / That god of war so glorious, deigned more than once / To look favorably at French theater. / … And if you want to judge people by their gold, / Know that theater brings profit not to be despised. / And your son in this sweet craft / Is better off than he was at home. / So stop repeating well-worn platitudes, / And instead rejoice at his genuine success. (Corneille / Hebanowski/Turdza, pp. 78–79)[4]

Clindor's father could only start rejoicing after going through a time of grief, coming close to taking his own life: "Farewell, my son is dead, so I wish to die" (ibid., p. 78). He had to appreciate the fact that his son was alive before he could accept the son's new profession.

Enchantment, entertainment, fun, illusion – for the power and culture elite, the "mighty" and the "commoners": Is this what theater's role boils down to? The "self-advertisement" practiced by Corneille's magician makes no mention of seeking truth, although in their pursuit of the truth the characters of French tragedy provoke, expose, torment themselves and others – suffice it to mention the plot episodes in Corneille's *Cinna* (1642), or in Racine's *Britannicus* (1669); even Molière's Baroque *Dom Juan* (1665) is a part of this trend, not to mention Orgon's family rebelling against Tartuffe's influence. Is this a sign of the Baroque's turning point leading to the neoclassical era with its invincible faith in the availability – through rules of art or laws of physics – of a rational order?

---

[4] Unless an English edition is listed in the references, English translations are based on the Polish text.

2. Two Messieurs T…,
or deception as madness[5]

It is precisely faith in rationality – as natural reason – that we find in *Tartuffe, or the Impostor* (*Tartuffe, ou l'Imposteur*, 1664–1669), the comedy that got Molière into trouble with the Company of the Holy Sacrament but which interests us here for reasons other than the writer's conflict at the time. Rachmiel Brandwajn underlines that the counterbalance to the hypocrite is not, as one might guess by analogy with *The Misanthrope* or *The Learned Ladies*, the sensible uncle, Elmire's brother; he, too, is tainted with cliquish thinking, in this case that typical of "the worldview of a lackey" (Brandwajn, 1968, pp. 64–68). Theater within the theater seems to be losing its innocence. If it is not in the spotlight itself, at least it suggests a different modus operandi: Its realm is expanded to include masquerade, which proceeds in time and in space, in the world and in the family. Perhaps even (and Tartuffe's monographer notes this kind of interpretation of the character), Tartuffe himself does not realize how far he is from the model of piety he is creating[6].

Erich Auerbach juxtaposes Molière's fraudster with the less ambiguous Onuphrius from La Bruyère's *Characters*. He does so out of concern for probability, something the comedic character supposedly contradicts – he underlines his own piety too openly, to excess, making mention of hair shirts and discipline, harshly encroaching upon the rights of his closest heirs. Onuphrius proceeds differently; he is the palace pharisee seeking acclaim and admiration, in his "contemplation" comparing himself quietly to the other courtiers as if to poor sinful tax collectors; he is

---

[5] Here I do not bring up the possibility of interpreting both plays from the perspective of Foucault's study *Folie et déraison. Histoire de la folie à l'âge classique*, Éd. Plon, 1961. The desacralization of madness, its isolation, possibly its treatment, applies to the lovers overwhelmed by mimetic desire in *The Argument* or the arrogant Europeans from Marivaux's *The Island of Reason*.

[6] Coquelin the Elder from (among others) the Comédie Française (Benoît Constant Coquelin, 1841–1909) apparently wrote in a brochure from 1884 that "Tartuffe is a mystic, Tartuffe believes… He experienced on himself how scruples should be dispelled. He distorted his conscience in a curious way; for the first time, the great fraudster was hoist with his own petard" (Brandwajn, ibid., pp. 84–85). True enough, in Scene 3 of Act III (the wooing of Elmire, Orgon's wife), Tartuffe's language undergoes a mystical alienation… unless this is in fact an intentional but ineffective manipulation.

pious in public, whereas piety to the moralist is a synonym of "sincere virtue" (*vertu sincère*), its antonym being a sanctimonious hypocrite (La Bruyère, 1965, pp. 345–347). In fact Molière sets out to expose not only the hypocrisy of *(faux) dévots*, but also the hidden intent of people pretending to be pious – to seize political power (and not just to "rule over people's hearts and minds"): If Molière clashed with their clique at the court, it was out of concern for the endangered rule of law, in which the arbiter should be the king broadcasting his "absolute" power and not any secret – real or mythical – society. But Tartuffe, as both his Polish translators sensed correctly, aims higher: He imitates *saints*, not "just" pious people. Thus, he delegitimizes not only the sociopolitical order but also the theological one. In the case of this character, aiming "higher" comes with being more distinctive – the needs of theater support this. Nevertheless, out of concern for the quality of literary comedy, Nicolas Boileau accused Molière of having overdone the expressiveness of his characters and mixing the lowbrow style of the farce with a comedy of manners (Auerbach, 1974, pp. 368–370). Indeed, Molière did not shun grotesque tones, raising to absurd levels the possibilities offered by his characters' monomania, the excess of one quality or inclination. "One can see in Molière's art the greatest measure of realism which could still please in the fully developed classical literature of the France of Louis XIV" (ibid., p. 370). This was the same classicism in which harmony, reason and nature[7] defined the boundaries of true, worthy art (ibid.), and whose social ideal was summarized in the stance of the *honnête homme*, positing as broad an education as possible, obliterating any traces of specialization out of concern for universalism, including that of language; what was not universal, became ridiculous (ibid., p. 372).

---

[7] Human "nature" was seen as something completely different in the 17th than in the 18th century, as Auerbach underlines in the same chapter: Nature was not considered in opposition to civilization and its education, nature was not associated solely with primeval cultures, folk inspirations or free space; nature, naturalness (*le naturel*) came from a good education that taught adaptation to any conditions, was associated with all that is rational and seemly (*bienséant*) and, above all, that "moves the human heart at all times and in all places" ("ce qui avait ému le cœur humain en tout temps et en tous lieux", ibid., p. 390). In this sense, the circumstances of asking nature about the first infidelity in Marivaux's *The Argument* illustrate the 17th-century paradigm: A good(?) education does not take away nature's right to speak. It was not until the second half of the century that Rousseau made a radical break with this thinking.

The belief in equilibrium guaranteed by constant human nature (ibid., p. 390) did not outlive Louis XIV himself. Even before the regency of Philip of Orleans, announced after Louis's death on September 1, 1715, proclaimed freedom of morals (at the top of the power ladder), on the fringes of the royal court and in Paris among financial circles there developed parallel, rival cultural circulations respecting the ethical façade guarded by the old king. The dichotomy between "being" and "appearing" (*être et paraître*) emerged and deepened. For over half a century the masquerade-like model of high society's life, also outside the carnival, would be reflected in comedies and novels, from Lesage and Marivaux to Crébillon fils and the Abbé Prévost. Inscribed into the ritual of play, momentary pleasure, charming detail and surprise, ultimately incorporated into the framework of the aesthetics known as Rococo – with its emphasis on irregularity and changeability, art from the close of the neoclassical period, before the arrival of the first naturalism (also referred to as primitivism or sentimentalism), also showed the dark, disturbing side of that changeable and unpredictable reality. Just like the ailing Louis XIV conducting treasury-depleting wars with variable success, the protagonist of Alain-René Lesage's comedy from 1709, Turcaret the financier, not only does not inspire admiration or trust but – as a victim of those smarter than him (truly or supposedly aristocratic scroungers) – actually deserves pity bordering on contempt, if not satisfaction.

Molière's audiences could experience catharsis when the royal officer came to arrest, instead of Orgon, Tartuffe the cynical plotter himself. At the end of *Turcaret, or the Financier*, it is his valet Frontin, having led the cheating Chevalier up the garden path, who takes the initiative because, contrary to the bankrupt Turcaret, he has capital at his disposal that the Baroness's disloyal servant Lisette agrees to share with him. As the eponymous hero falls into his cheated partners' hands, Frontin announces: "Monsieur Turcaret's reign is over. Mine will begin" (Lesage, 1951, act V, scene 18 – in the French edition: scene 14).

Here is France 40 years after *Tartuffe*: The family we knew from Orgon's household (grandmother, father, children, virtuous wife) no longer exists. The financier and his wife having been separated for ten years, Turcaret the indefatigable admirer of "belles parisiennes" introduces himself as a widower while paying his spouse to stay put in her provincial abode, from whence

– in the face of some overdue payments – she sets off anyway to conquer Paris, where she falls victim to the self-interested courtship of the bon vivant Marquis. Love without capital – or at least a view to some – cannot survive; it sometimes involves triangular relationships, which can include multi-layered fraud: Turcaret courts the Baroness, who is responsive to the wooing of the Chevalier, who plans to get her to help him fleece the Financier; everyone is led up the garden path by Frontin and his accomplice, Lisette the maid. The only positive character in Lesage's comedy (?) (besides Marine, whom the Baroness fires for frankness and honesty already in act I – Lesage, 1951, p. 25), i.e. the only person constant in their feelings, is Madame Jacob, Turcaret's sister. She makes a living for herself, her husband and children by peddling cosmetics and… matchmaking. It is worth noting that when she hears of her brother's arrest, she rushes to him, as does his wife. However, whereas Madame Turcaret wants "to revile him: I feel I am his wife!" (ibid., V, 15), Madame Jacob declares: "Despite his base acts, I feel sorry for him! I shall use all my influence to save him, I feel I am his sister" (ibid., V, 14).

What threats hang over this society of appearances, chasing after money? Louis XIV himself secured its reign when he decided to fill official positions with candidates legitimized by their money pouches rather than merit and/or birth, as used to be the case. The pessimistic tone of Lesage's play stems from its verdict regarding lies: In the end, they are rewarded, not punished; the only person to be punished is Turcaret, but not because he lied but because he was inept at plotting. It is the same with the Baroness and the Chevalier, who manage to get themselves out of trouble at little cost (the loss of new profits). You could say that in the face of disintegrating social ethics, in Lesage's plays neoclassicism as a style of cleansing by presenting the truth reached the boundary of catharsis, facing a similar dilemma as the French novel did prior to the publication of Rousseau's *New Heloise* in 1761: What should one choose – likelihood (literary truth) or a moral message? (May, 1963, p. 47). It had become impossible to defend one without weakening the other. The only possible solution was to seek a new perspective. Lesage, who was in conflict with the company of the Comédie Française, went on to find his place in the aesthetics of fairground comic opera (*l'opéra-comique de la foire*). This theater, redefining itself in

the battle against restrictions imposed by its rivals, the Comédie Française and later also the Comédie Italienne, promoted burlesque and spontaneity woven into a genre or fairy-tale convention, or sometimes one that was allegorical or parodic. The conventionality of the characters, invoking *commedia dell'arte* or the topsy-turvy world model, only highlighted their naturalness concentrated in language. The ubiquitous masks became a sign that there must be truth concealed beneath them, a truth either pleasant or cruel; and the other way round: to prevail, truth must join in the masquerade. The fairground convention making things unreal toned down the harshness of the diagnosis being offered.

A different approach to this challenge was chosen by Marivaux (1688–1763), who overcame neoclassical conventions through empirical philosophy combined with the naturalistic though slightly conventionalized aesthetics of Italian theater.

3. Morality play of impossibility: from Marivaux to Diderot

In the oeuvre of the greatest 18th-century French playwright, the work that speaks the most explicitly – besides the one-act *Les Sincères* (*A Case of Sincerity* – 1739, Comédie Italienne) – about the relationship between truth and illusion, which is morally condemned as falsehood, is his last known play: the one-act *Les Acteurs de bonne foi* (*The Actors in Good Faith*) published in 1757 in *Conservateur* magazine. The original title itself already contains a paradox: involuntary actors? Sincere, genuine actors? They play themselves, like the servants in the first of the inner plays (*comédie des valets*), according to the idea of the scriptwriter – Merlin the servant, commissioned to write it by Éraste who wants to honor his aunt, Madame Amelin, and to celebrate his own betrothal to Angélique. But then there appears – expected by no one and with parts assigned without warning to some of those involved (Éraste and Angélique with her mother) – a "comedy of masters" (*comédie des maîtres*), Madame Amelin's revenge on the fiancée's mother who is against the first comedy being staged at her home. Because the fact is, that innocent play is not without its controversy. It was to have shown the love of two out of four lovers being put to the test, to the satisfaction of the other two. Instead of pleasure, however, jealousy appeared, its cause being

summarized the most clearly by Blaise when he complains to his mistress about Colette's (supposed?) infidelity:

> … madam, this wretched play mocks me personally; Colette feigns affection for monsieur Merlin, monsieur Merlin pretends to reciprocate; and although it is a comedy, it's all true, madam; because *they are only pretending to pretend so as to fool us better* [emphasis mine – IZ], and actually do love each other on the sly … (English translation based on the Polish literary translation, scene 12)

If we invoke Wielechowska's analysis of Girard once more, we might notice the elite audience's symbolic violence toward naïve servants unfamiliar with the subtle difference between fiction and games on one hand and truth on the other, like those Athenian craftsmen in Shakespeare. But Marivaux develops the metatheatrical experiment into a game in which unaware actors – Éraste and Angélique – are subjected to a test even more dramatic than the one from the "comedy of servants" which already had unpleasant consequences for at least half of the participants; for the two who are betrothed, the test is a test of faithfulness not in the face of temptations but in the face of violence from family authority and money. An enemy of such violence, Marivaux considers it conditionally, promptly revoking the danger hanging over the young couple as soon as the Solicitor reads out the prenuptial agreement: Éraste will not have to marry his aunt's forty-year-old friend, as the two ladies were only having a laugh at his expense to teach Madame Argante a lesson. But the test to which his aunt put him has shown Éraste to be unable to defend his relationship with Angélique: His sighs and cries remain powerless. What hope does this hold for their relationship in future? Exposing the fiancé's weakness in the game improvised by the fiendish aunt, even if the contract stayed unchanged, revealed the true balance of power in the family. Frédéric Deloffre saw the "test" as one of the structural models of his comedies. It is a paradoxical test: It often costs its author just as much as the beloved he or she is testing. This is the case in the master playwright's earlier plays, e.g. *The Test* (*Épreuve*) from 1740.

Witness to an ethical split between appearances and inner life, Marivaux also uncovered the ambiguity of theatrical conventions: It is hard to separate the art of illusion (accepted, autotelic, safe, trust-inspiring) from lying (self-seeking, stemming from bad faith, assuming someone else's harm, dangerous – especially when it is

masked as truth and stops arousing mistrust). But this means that the pretenders can lose their awareness that they wear masks. In such a case, how do you make them responsible for the hoax? The solution would be the theatralization of the scene: turning the game of illusion into the subject; this is a method allowing the truth to be sought – over and above the immediacy of the masquerade. The latter is also thematized in language, in the famous *marivaudage*, in a dual register composed of a conscious layer permeated with signs of unconscious or unintentional expression of repressed desire, resentment, fear – almost two centuries before Freud.

Diderot understood metatheater differently. In treatises accompanying his two drama experiments from 1757 and 1758 (*The Natural Son*, *The Father of the Family*) – in *Conversations on…* and in *Discours sur la poésie dramatique* – he outlined a new aesthetics enabling theater to fulfill its moral mission. Interestingly, the characters in the first of these plays are only meant to act out an event they actually experienced – as a form of celebration, but above all to instill it into themselves as a lesson in virtue through a renewing imitation of their own magnanimity. Placing nature (naturalness) before reason, Diderot underlined the value of nonverbal truth: Strong feelings make people unable to speak or make their speech chaotic, inarticulate. This is nothing to be afraid of, it is where the truth of the situation defining the truth of the character lies. It is extraordinary that Diderot's demand-making theater seemed not to have heard the warning coming from Marivaux's metatheatrical remarks: The new plays (still called comedies) assumed – like Rousseau's *New Heloise* on which he was working during this time – the complete conscious mutual transparency of characters (Dorval's unexpected recognition of Lysimond as his long-lost father, who will make him realize that his beloved Rosalie is in fact his sister, only confirms this rule). In a paradox (?), Diderot the atheist sacralizes art, assuming – for the first time since the Middle Ages (?) – a communion of sensations and feelings (shared with the characters) existing between the stage and the audience (Diderot/ Dębowski, 2008, pp. 77–78). However, Diderot's plays got a cool reception in Paris. Contrary to the writer's belief, the "spirit of the times" had not matured in audiences (ibid., pp. 83–84). It took Lessing's determination to finally translate aesthetical theorems into living theatrical practice. And the recommendation that characters be built from the same fabric that audiences are made of, was also the main demand in *The Hamburg Dramaturgy* (1767–1769).

The "serious genre" of drama that Diderot proposed in *Discours sur la poésie dramatique* – intermediate between "serious comedy" and "domestic tragedy" – corresponded with the earlier experimental ideas of Lessing, Diderot's translator in 1760 but also the author of *Miss Sara Sampson* published five years earlier, a play in which – with the help of less perfect characters – he showed similar dilemmas (the need to renounce one's love) as Dorval and Rosalie experience in *The Natural Son*. Diderot's demand for truth was more of a demand for authenticity: The father of the Encyclopédie did not believe in a metaphysical absolute, but he was interested in the accurate reproduction of immanent feelings and relationships, as he had already posited in the words of Mirzoza, the sultan's mistress in *The Indiscreet Jewels* (1751): "I also know that the perfection of a dramatic piece consists in the exact imitation of an action, so that the spectator, continually deceived, imagines he is present at the very action" (Immer, Müller, 2015, p. 138). Lessing would need Diderot in 1760 to question the perspective of classical tragedy being solely entitled to discuss serious issues. Lessing used Diderot's authority as an art critic to shield himself from the authority of Gottsched, a defender of tragedy deposed by its younger sister, the domestic or bourgeois tragedy (ibid., p. 139).

4. The "Romantic lie" and its two facets

Friedrich Schiller was 25 when he celebrated the success of *Intrigue and Love* (*Kabale und Liebe*) in 1784[8]. This play concluding the "Sturm und Drang" (Storm and Stress) period was severely criticized by Auerbach in his essay "Miller the Musician." According to the critic, a melodramatic structure could not carry a revolutionary message[9]. The end result was a series of black-and-white simplifications that were ultimately comical, and some bombastic antics by the characters – Ferdinand, Lady Milford and Luise herself (Auerbach, 1974).

---

[8] The original title *Luise Millerin* has been preserved in the opera by Verdi, produced in spring 2018 by the Metropolitan Opera and featuring Piotr Beczała as Ferdinand/Rodolfo (the librettist changed the romantic lead's name). Schiller changed the title at the advice of his friend, Iffland (Schiller, 1976, pp. XXXII–XXXIII).

[9] "*Luise Millerin* is much more a political and even a demagogic play than a truly realistic one" (Auerbach, 1957, p. 388).

If we read Schiller's play with the help of René Girard's formula, which he renewed in the monograph *Mensonge romantique et vérité romanesque* (*Deceit, Desire and the Novel*), we can look at Ferdinand and Luise not so much as victims of a political intrigue but as victims of the "Romantic lie" that engulfed the imagination (and sensitivity) of the egotistic Ferdinand, son of the cynical and ambitious President, the prince's highest official. "These individualisms professed with fanfare merely hide a new form of imitation. Romantic revulsions … usually conceal a morbid concern for the Other" (Girard, 1996, p. 43). That Other becomes an object of hatred: "someone who prevents us from satisfying a desire which he himself has inspired in us" (ibid., p. 40). Is this how we can perceive Luise? In the first act Ferdinand declares: "I see clearly every emotion on your face. … If only this mirror stays unsullied, there will be no more clouds in the sky" (Schiller, p. 18). He has faith in the extraordinariness of his love, but the guarantor (guardian?) and, as it will transpire, the obstacle or even destructor of this highest good in which he invests his entire (excessively loved) self, in Ferdinand's case is "a nothingness called Luise," as the girl introduces herself (ibid., p. 19). "Is my escutcheon worth more than the judgment of heaven in my Luise's eyes?" Ferdinand asks rhetorically (ibid.). As the mediator of the "triangular desire" to which her "boy" (as he calls himself) seems to surrender rightly points out: "Do you not feel that your hopes rend my heart into pieces like fiends?" (ibid, p. 20). Luise shares her forebodings with him, whereas Ferdinand's dreams include the unintentional anticipation of the final tragic scene between the lovers, i.e. Luise's poisoning and the suicide of her lover who accused her of infidelity: "for thee I shall receive every wound"; from the point of view of the ending, this protestation assumes a somberly ironic tone: Ferdinand will drink the beverage he poisoned, which he had earlier asked to be made for the girl.

Can the excess of pathos in Ferdinand's words be ascribed to the writer's lack of creative maturity (Auerbach, 1974)? "The romantic *vaniteux* … convinces himself that he is thoroughly original" (Girard, 2001, p. 44). Ferdinand is not meant to be a hero without blemish. His excessive eloquence is in fact a sign of alienation: He loves the power and extraordinariness of his love for Luise who discarded convention; but does he love Luise? Luise tries to measure up to the role he has assigned her, but

draws back terrified. She will step on the path toward death (but not infidelity) only because – as she believes – she is thus saving her father imprisoned by the prince. The uniqueness of *Intrigue and Love* – as Auerbach admits – lies in the fact that, for the first time, the fate of an individual incorporates the quintessence of a situation being experienced by an entire community (Europe in the shadow of the French Revolution, the German provinces). Standing opposite the tragic lover lost in his loftiness is the honest and sensitive daughter of a music teacher, defending the truth of love that Schiller and his colleagues took over from Rousseau: love that is tragic and noble, natural and directly accessible to everyone. "Its simplest and purest form appeared to be a condition of natural virtue, and its freedom in the face of mere convention was considered an inalienable natural right" (Auerbach, 1957, p. 389).

There is a similarity between the situation of Schiller's Luise and Rosette from Alfred de Musset's comedy *No Trifling with Love*: Young master Perdican, thinking himself rejected by his cousin Camille whom his father the Baron wants him to marry, toys with the feelings of a "peasant girl" ultimately using her as "bait" for his own ego and Camille's jealousy; it works, because the latter decides not to return to the convent, finally telling Perdican that she loves him. This happens in the chapel, where Rosette also comes, uninvited and unnoticed. Her cry in response to the double ecstatic confession of Camille and Perdican will, however, remain a symbol of mystery: Insofar as the two cousins discover their lies and identify them as resulting from Camille's ambition and Perdican's vanity, neither they nor the audience ever find out how Rosette died. The only certain thing is what Camille announces to her beloved at the end: "She is dead. Farewell, Perdican" (Musset, 1953, p. 88 – act III, scene 8).

"Innocence designated from the start as a guilt offering" for the egoisms of the two protagonists (Musset/Canal, 1996, p. 114) who are social and intellectual equals, Rosette comes close to the role of the mediator from Girard's theory: Both Perdican and Camille take her as an example of ideal, honest (because unthinking!) love (ibid., pp. 67–69, 75–78 – act III, scene 3 and 6); their own behavior toward Rosette, however, is callous: They continue their dispute instead of attending to the unconscious girl who, after all, is Camille's foster sister (ibid., pp. 78–79). Their thoughtless cruelty is an ominous indication of the finale.

If, following *Encyklopedia teatru polskiego* [Encyclopedia of Polish Theater], we acknowledge the special role of metatheater[10] in building meaning that "undergoes splitting – from referencing the external, it shifts to the level of self-reference: commentary on the process or means of creating theater," the striking thing about the last two plays discussed above – Schiller's and Musset's – is how they move away from the structures of theater within the theater or theater referencing itself. Their metatheater does not involve seeking truth by the staging of illusions, they forgo spying on frauds and psycho-dramatic imitation of themselves by characters who trust – rightly or not – their own moral strength. Just like in the theater of the world, the image appearing in this theater is that of a topsy-turvy world, Perdican and Camille play out the scene from *Hamlet* in a topsy-turvy manner: At first the cousin advises his beloved against going to the convent when she asks for his opinion. The scene itself smells of provocation from the maiden in love (as yet unconsciously) with Perdican; it is only during the scene's final retort that Perdican, pushed to the limits by his cousin's mockery and sarcasm, offers a pastiche of the Shakespearian model: "Farewell, Camille. Return to the convent …" (Musset, 1953, pp. 43–57 – act II, scene 5). This reversal concerns not only the perception of one's beloved (after all, Marivaux's Blaise and Lisette could also be jealous for no reason) but, above all, the contaminated nature of love or, more precisely, yearning for love, desire for another person, toward whom Romantic heroes adopt an intentional attitude (as they would toward things) rather than a dialogical one (as they should toward another human being) – if we espouse Father Józef Tischner's differentiation (in which he speaks of 'an opening up' and not 'an attitude', making it more dynamic) which is immersed in Martin Buber's philosophy of dialogue (Tischner, 2001, pp. 9–10). The (female) Other functions as a sign – hidden in bushes or behind a drape in Musset, forced into silence and insulted despite insisting on her innocence in Schiller – as a stimulus and an obstacle, prize and punishment all in one, on the way to the male character's fulfillment. "If desire turns us toward good, then thinking

---

[10] "Metatheater – A play or performance self-reflective in character, whose development illustrates issues related to theater. The metatheatrical aspect can be designed by the playwright within the text's structure or added by the director in the staging process" (ibid.).

turns us toward truth" (ibid., p. 46). What has happened to the world that such a simple mapping has become impossible?

In the drama of existence, which theater tries to play out in its own way and in every time, Father Tischner defines a significant difference between a lie and an illusion:

Lying assumes knowledge of the truth: The liar knows how things really are, but says they are otherwise – thus lying. In illusion the situation is different – the speaker does not know how things really are, to him they seem to be as he says and hears them, so he is convinced he is speaking the truth (ibid., 143).

Romantic drama superimposes lies and illusions on each other: Luise lies because she has to; however, to Ferdinand the illusion of infidelity created by her lie is the truth, but this is because his desire becomes immersed in itself instead of taking note of the nonverbal signs that his beloved, committed to silence by a sacred vow, gives him. Are Perdican and Camille lying, or are they victims of their own illusion? To Rosette, their conduct in the final scene in the chapel is unequivocal: She hears (and sees?) that she has been cheated.

What we see in the boundless individualism of the Romantics is not only the release of stress from centuries of practicing good manners under the guidance and laws of the authorities of the Ancien Régime; if the Romantic hero declares war on evil – in politics or morals – does it mean he desires good? The condition of an ethical stance on the stage of *theatrum mundi*, according to Tischner, is acknowledgement of the truth: "My self is infinitely responsible before my fellow human," and therefore responsible for the whole world, because "you are either responsible for everything or for nothing. … Between the stance of Abel and Cain there is no other option" (ibid., p. 50). Instead of having a choice, we are left to wander confused.

References

Aristotle (1989). *Poetyka* (Polish edition of *Poetics*, transl. H. Podbielski). Wrocław: Ossolineum.

Auerbach, E. (1968). *Mimesis: rzeczywistość przedstawiona w literaturze Zachodu* (Polish edition of *Mimesis: Dargestellte Wirklichkeit in der abendländischen Literatur*, transl. Z. Żabicki). Warszawa: Państwowy Instytut Wydawniczy. [English quotes in the text are taken from: Auerbach, E. (1957). *Mimesis: The Representation of Reality in Western Literature*. Princeton, NJ: Princeton University Press].

Boileau, N. (1989). *Sztuka poetycka* (Polish edition of *L'art poétique*, transl. M. Grzędzielska). Lublin: Wydawnictwo UMCS.

Brandwajn, R. (1965). *Twarz i maska. Rzecz o Świętoszku Moliera* [Faces and Masks. On Moliere's Tartuffe]. Warszawa: Wiedza Powszechna.

Brandwajn, R. (1968). *Świętoszek (Tartuffe) Moliera* [Moliere's Tartuffe]. Warszawa: Państwowe Zakłady Wydawnictw Szkolnych.

Corneille, P. (1993). *Iluzja komiczna* (Polish edition of *L'Illusion comique*, transl. S. Hebanowski & W. Turdza). Typescript.

Deloffre, F. (1955). *Une préciosité nouvelle. Marivaux et le marivaudage. Étude de langue et de style*. Paris: Les Belles Lettres.

Diderot, D. (2008). *Pisma estetyczno-teatralne* (Polish editions of Diderot's works on aesthetics and theater, M. Dębowski, Ed., transl. M. Dębowski, J. Kott, E. Rzadkowska, & A. Siemiek). Gdańsk: słowo/obraz terytoria.

Dygul, J. (2010). Komedia w komedii, czyli o metateatralności w siedemnastowiecznej dramaturgii włoskiej [Comedy within the comedy, or on metatheatricality in 17th-century Italian drama]. *Odrodzenie i Reformacja w Polsce*, *54*, 41–67.

Forestier, G. (1981/1996). *Le Théâtre dans le théâtre sur la scène française du XVIIe siècle*. Geneve: Droz.

Foucault, M. (1961). *Folie et déraison. Histoire de la folie à l'âge classique*. Paris: Plon.

Girard, R. (2001). *Prawda powieściowa i kłamstwo romantyczne* (Polish edition of *Mensonge romantique et vérité romanesque*, transl. K. Kot). Warszawa: KR. [English quotes in the text are taken from: René Girard (1965), *Deceit, Desire and the Novel. Self and Other in Literary Structure* (transl. Y. Freccero). Baltimore: The Johns Hopkins Press].

Goethe, J. W. (1951). Goetz von Berlichingen (French transl. A. Stapfer). In: *Théâtre complet*. Paris: Librairie Gallimard.

Goethe, J. W. (1957). *Z mojego życia; zmyślenie i prawda* (Polish edition of *Aus meinem Leben: Dichtung und Wahrheit*, transl. A. Guttry). Warszawa: Państwowy Instytut Wydawniczy.

Haac, O. A. (1973). *Marivaux*. New York: Twayne Publishers, Inc.

Immer, N., Müller, O. (2015). Le Diderot de Lessing : de « douces larmes » pour servir à la purification du goût national. *Recherches sur Diderot et sur l'Encyclopédie*, *50*, 121–140.

Kowalski, W. (2018). Recenzje: Szekspir w rozkroku [Reviews: Shakespeare Astraddle]. Retrieved from https://teatrdlawszystkich.eu/szekspir-w-rozkroku/

Kowzan, T. (1976). L'Art en abyme. *Diogène*, 96, 67–92.

La Bruyère, J. (1965). *Les Caractères de Théophraste traduits du grec avec Les Caractères ou Les Moeurs de ce siècle, chronologie et préface par R. Pignarre*. Paris: Garnier-Flammarion.

Lesage, A.-R. (1970). *Turcaret* (B. Blanc, Ed.). Paris: Librairie Larousse.

Lesage, A.-R. (1951). *Turcaret, czyli finansista. Komedia w pięciu aktach* (Polish edition of *Turcaret ou le Financier*, A. Jakubiszyn, Ed., transl. T. Żeleński [Boy]). Wrocław: Ossolineum.

Marivaux, P. de (n/a). *Aktorzy mimo woli* (Polish edition of *Les Acteurs de bonne* foi, transl. I. Zatorska). Own printout.

Marivaux, P. de (1999). *Spór* (Polish edition of *La Dispute*, transl. B. Skoczyński). In: M. Ziębiński (Director), *Teatr telewizji*. Warszawa: Telewizja Polska.

Marivaux, P. de (1989/1996). *Théâtre complet* (Vol. 1, F. Deloffre & F. Rubellin, Eds.). Paris: Bordas/Dunond.

Marivaux, P. de (1992). *Théâtre complet* (F. Deloffre & F. Rubellin, Eds.). Paris: Bordas.

May, G. (1963). *Le dilemme du roman au XVIIIe siècle : étude sur les rapports du roman et de la critique, 1715–1761*. Paris: Presses Universitaires de France.

Mitosek, Z. (1992). Wprowadzenie [Introduction]. In: Z. Mitosek (Ed.), *Mimesis w literaturze, kulturze i sztuce* [Mimesis in Literature, Culture, and Art] (pp. 5–32). Warszawa: Państwowe Wydawnictwo Naukowe.

Molier (1883). *Świętoszek, czyli szalbierz* (Polish edition of *Tartuffe, ou l'Imposteur*, transl. A. Urbański). Lwów: Księgarnia Polska.

Molier (1987). *Świętoszek* (Polish edition of *Tartuffe*, transl. T. Żeleński [Boy], Ed.). Łódź: Wydawnictwo Łódzkie. Retrieved from http://lektury.crib. pl/swietoszek-molier-tekst-utworu.html or https://pl.wikisource.org/wiki/ Dyskusja:Świętoszek

Molière (1875). *Świętoszek. Tartuffe: komedya w pięciu aktach wierszem* (Polish edition of *Tartuffe*, transl. K. Zalewski). Warszawa: S. Lewental.

Musset, A. de (1996). *On ne badine pas avec l'amour. Proverbe* (D. A. Canal, Ed.). Paris: Larousse-Bordas.

Musset, A. de (1953). *Nie igra się z miłością. Komedia w trzech aktach* (Polish edition of *On ne badine pas avec l'amour*, L. Łopatyńska, Ed., transl. T. Żeleński [Boy]). Wrocław: Ossolineum.

Palus, B. (2017). *Literatura in statu nascendi. Problemy estetyczne i genologiczne prozy literackiej Denisa Diderota* [Literature in statu nascendi. Ethical and Genological Issues in the Literary Prose of Denis Diderot]. Wrocław: Oficyna Wydawnicza ATUT.

Rousset, J. (1962). *Forme et signification. Essais sur les structures littéraires de Corneille à Claudel*. Paris: J. Corti.

Schiller, F. (1976). *Intryga i miłość* (Polish edition of *Kabale und Liebe*, O. Dobijanka-Witczakowa, Ed., transl. A. M. Swinarski). Wrocław: Ossolineum.

Shakespeare, W. (1992). *Sen nocy letniej. Kupiec wenecki* (Polish editions of *A Midsummer Night's Dream* and *The Merchant of Venice*, A. Jankowski, Ed., transl. S. Barańczak). Poznań: Wydawnictwo W Drodze.

Tischner, ks. J. (2001). *Filozofia dramatu. Wprowadzenie* [Philosophy of Drama: Introduction]. Kraków: Wydawnictwo Znak.

Van Thiegem, Ph. (1971). *Główne doktryny literackie we Francji. Od Plejady do surrealizmu* (Polish edition of *Petite histoire des grandes doctrines littéraires en France: De la Pléiade au surréalisme*, transl. M. Wodzyńska-Walicka & E. Maszewska). Warszawa: Państwowy Instytut Wydawniczy.

Wąchocka, E. Metateatr [Metatheater]. In: *Encyklopedia teatru polskiego* [Encyclopedia of Polish Theater]. Retrieved from http://encyklopediateatru. pl/hasla/108/metateatr

Wielechowska, K. (2013). Cicha wiedza Hippolity – Sen nocy letniej w lekturze René Girarda [Hippolyta's tacit knowledge: *A Midsummer Night's Dream* in the interpretation of René Girard]. *Roczniki Kulturoznawcze*, *4*(2), 97–122.

Wojciech Sajkowski
ORCID 0000-0002-8585-4149
Adam Mickiewicz University, Poznań

Honesty as a Trait of Non-Civilized Man in the French Image of Southern Slavs at the Turn of the 18th and 19th Centuries[1]

Abstract

The paper deals with the French view on the morality of peoples considered uncivilized, specifically Southern Slavs inhabiting the territories of today's Croatia and Montenegro. The problem of morality is analyzed first of all in the context of natural honesty, which was attributed to peoples considered uncivilized. The analysis uses not only the testimonies of French authors, but also French-language descriptions that were popular in France during the Enlightenment, especially the works of Alberto Fortis and Stefano Zannowich. These sources shaping the image of the Morlachs (Slavic shepherds living in provincial Dalmatia) or Montenegrins, which were popular in the second half of the 18th century, are compared with documents related to French rule in Dalmatia (1806–1813). The analysis has allowed to show what place the Slavic communities living in the peripheries of Europe at the time had in the Enlightenment's discussions on the morality of civilized and savage people.

Keywords: Morlachs, Montenegrins, sincerity, falseness, Enlightenment, Dalmatia, Illyrian Provinces

The issue of dishonesty, an inclination toward confabulation, and, on the other hand, truthfulness, honesty, or simply the inability to produce untrue assertions, is an extremely important aspect of descriptions of the character of peoples, nationalities, or any other social communities in modern times. In the second half of the 17th and at the beginning of the 18th century, the problem of honesty became one of the most important aspects of

---

the discussion about the possibility of practicing Christian morality in the corrupt world of the European elites. In the second half of the 17th century many authors and moralists were discussing the possibility of merging Christian morality with the practice of social life[2]. And although in the second half of the 18th century philosophers of the Enlightenment strongly emphasized the social legitimization of virtues, and the necessity of practicing them in the context of social rituals, the opinion that civilization (along with its most sophisticated forms) was a factor which may lead to the corruption of morals was still powerful. The difference from earlier criticism was the change of reference point. Now, the artificiality of worldly manners was not juxtaposed with evangelical virtues, but with the natural morality of communities considered uncivilized. Among the nations considered less civilized, one could point not only to those living in distant and exotic parts of the world, but also to those living in some remote European regions, such as the highlanders from Valais described by Jean Jacques Rousseau in *The New Heloise*. Hence, criticism of the corruption of the elites' morality was offered not in the context of religion but in the perspective of the "natural history" of mankind, i.e. anthropology[3]. The recognition of peoples considered uncivilized as a model of truthfulness or sincerity includes the ideal of the "good savage," which became an extremely popular motif present in scientific and literary works of the 18th century. "Good savage" (*bon sauvage*) or "noble savage" are terms that define an idealized image of primitive man who, living away from civilization, had to preserve certain moral virtues, such as living in harmony with nature, innocence, disinterestedness, but also sincerity. The term *bon sauvage* was first used by French writer and traveler Marc Lescarbot in 1609, but it was Rousseau who

---

[2] This problem appears e.g. in Molière's *Misanthrope* and La Bruyère's *Characters*. The theoretical foundation for the possibility of combining virtue with high society life was provided by the writings of Malebranche, which were opposed to Pascal's strict morality (Bury, 1996, pp. 123–125, 151, 199–202.).

[3] Already in 1724 Joseph-François Lafitau delivered the first example of comparative anthropology (Motsch, 2001, pp. 3–6 ; Racault, 2003, pp. 17–22), followed a few decades later by the monumental work *Histoire naturelle de l'homme* by Georges-Louis Leclerc de Buffon, which also listed numerous examples from travel literature and on this basis divided mankind into uncivilized and civilized communities. Also Diderot, Helvetius, Voltaire, and Rousseau, based on their reading of travelogues, tried to determine the laws governing the development of humanity (Duchet, 1971, pp. 235–249).

was responsible for its popularization. Thanks to his work, the idea of the good savage became a popular motif of Enlightenment art and literature (Ellingson, 2001, p. XVI, 2).

The present paper sets out to show how the morals of societies considered uncivilized and the mores of populations living in the European peripheries of civilization were regarded as being connected with each other. French-language descriptions of Morlachs (Slavic shepherds living in provincial Dalmatia[4]) and Montenegrins will be analyzed in this context of the 18th-century discussions on natural morality.

Already in the second half of the 17th century, French literature mentioned the Southern Slavs outside the historical context of their barbaric origin, using a kind of pre-anthropological perspective. Commenting on the travelogue of François Pyrard, the first French traveler to reach India, geographer Pierre Duval, nephew of famous cartographer Nicolas Sanson (Petto, 2007, p. 149), juxtaposed the pirates from Malabar, who were said to be bold but cruel soldiers with, among others, the Morlachs (Duval, Pyrard 1679, p. 114). Duval compared many different groups of people from all over the globe, which – in his opinion – shared a similar way of life. This long list opened with the Cherokees and closed with the Pyrenean highlanders. Based on the same principle, Paul Hay du Chastelet called Morlachs "the Dalmatian Iroquois" in his *Politique Civile et Militaire des Venitiens* published in 1670 (Chastelet, 1670, p. 103). In the travelogue of Alberto Fortis published more than 100 years later, the primitiveness of the Morlachs was emphasized by their comparison to Tatars and North American Indians (Fortis, 1778, p. 67).

The latter of the above-mentioned authors is particularly important, since his travelogue entitled *Viaggio in Dalmazia* became the primary source of information about the Southern Slavs in Europe during the Enlightenment. Fortis was a naturalist who undertook a special mission for the Venetian Republic and tried to describe Dalmatia, with particular emphasis on its

---

[4] From the 16th century the ethnonym "Morlachs," coming from the Greek term "Black Vlachs" (*Mavro Vlasi*), was used not only to describe Romance-speaking Vlach shepherds but also Slavs, and even other groups which merged with them. This was possible because Slavs often became shepherds, and the Venetians began to use the name "Morlachs" (*Morlacchi*) for all people who were herdsmen, regardless of the language they used or the religion they practiced (Rapacka, 1997, p. 218).

natural resources which could later be used by Venice (Wolff, 2001, p. 2). Fortis's work became especially popular in France and determined the image of the Southern Slavs for a few decades (Sajkowski, 2013, p. 166–167).

Fortis was the first author to provide a detailed description of the customs of the people described as Morlachs, but what is even more important in the context of the present paper's focus, he tried to replace the previous negative image of Morlachian barbarity with a new portrayal containing elements of the ideal of the "good savage." In a letter addressed to English nobleman Lord Bute, he emphasized:

> You have, no doubt, often heard the Morlacchi described as a race of men, fierce, unreasonable, void of humanity, and capable of any crime. The inhabitants of the sea coast of Dalmatia tell many frightful stories about the cruelty of those people, that, induced by the avidity of plunder, they often proceeded to the most atrocious excesses of violence, by fire and sword. But these facts (though the truth of them is sufficiently authenticated, by the known veracity of those who relate them) are either of ancient date, or if some have happened in later times, they ought rather, from the characters they bear, to be ascribed to the corruption of a few individuals, than to the bad disposition of the nation in general. It is but too true, that, after the late wars with the Turks, the Morlacchi, habituated to murder and plunder with impunity, gave some examples. of cruelty and rapine: but what instance can be given of troops just returned from war, and dismissed from the exercise of arms, against the enemy of their sovereign, that have not peopled the woods and highways with thieves and assassins? I think it, however, a duty incumbent on me, to write what I personally saw relative to their customs, and inclinations, and thereby to form some apology for that nation, by which I was so well received, and treated with so much humanity (Fortis, 1778, p. 44).

In Fortis's travelogue the natural goodness of the Morlachs went hand in hand with sincerity, which often made them vulnerable to being exploited. Because of that, Italians who traded in Dalmatia could take advantage of the gullibility of these "good people," as Fortis called the Morlachs. For this reason, their trust diminished appreciably due to suspicion and fear of being deceived:

> The sincerity, trust, and honesty of these poor people, not only in contracts, but in all the ordinary actions of their life, would be

called simplicity and weakness among us. It is true, that the Italians, who trade in Dalmatia, and the littoral inhabitants themselves, have but too often taken advantage of this integrity (Fortis, 1778, p. 53).

According to Fortis, the savages were guided primarily by emotions, and their innate naivety went hand in hand with superstition, which in the opinion of philosophers of the Enlightenment was one of the greatest enemies of intelligent thinking. In 18th-century descriptions of Indians and indigenous Africans, superstitions were discussed in detail. Following these models, Fortis decided to describe Morlachian belief in vampires, werewolves, ghosts, and witchcraft. Ignorance and superstition were presented as the result of the credulity and naivety of the Morlachs, who did not know falseness and accepted the teachings of backward and obscurantist clergy without criticism. The latter were meant to have aroused religious fanaticism in the gentle hearts of the Morlachs (Fortis, 1778, p. 61–64).

Another testimony from which French readers could learn about the natural sincerity of the Southern Slavs was produced by Stefano Zannowich (Stefan Zanović; Stiépan Zannovitch), false prince of Albania, an impostor who, justifying his nobility, wrote books that invoked Albania and Southeastern Europe, most often in the spirit of exoticism and *turquerie* (Watzlawick, 1999, pp. 14–15; Vidan, 2004, p. 136). Perhaps Zannowich's most important work – as regards describing the Balkan peoples – is also the most special in his bibliography. It is a work dedicated to another adventurer, Šćepan Mali (Stephen the Little), who reigned in Montenegro from 1767 to 1773 (Mortier, 2000, pp. 44–45). This impostor of Slavic origin appeared in Montenegro in 1766. He pretended to be Tsar Peter III, the ex-husband of Tsarina Catherine II, and unified Montenegro, which at the time was divided as a result of quarrels among various Montenegrin families. The presentation of the case of Šćepan Mali necessitated a detailed description of Montenegro and its inhabitants, because French knowledge on this country was almost nil. Zannowich, as someone who wanted to appear as an external observer of European reality, often criticized the manners of European high society. In the story of Šćepan Mali, he referred to the presumed corruption of European elites and contrasted it with the customs of Montenegrin women, similarly to Rousseau who had compared the women of Paris and Valais in *The New Heloise*, or like Fortis

who had glorified the straightforwardness and innocence of the Morlachs. According to Zannowich, the Montenegrins, even if they were belligerent and had shown barbaric cruelty, also shared some common traits of character with good savages. "The purity of morals is much higher among Montenegrins than among civilized peoples,"[5] stated Zannowich (Zannovitch, 1784, pp. 1–2). Just like all savage peoples, including the Morlachs described by Fortis, Montenegrins were very hospitable. Although Turkish, Venetian or Ragusan incomers had to watch out, Slavs who appeared in their country could count on a friendly reception. Other positive qualities of the Montenegrin character also duplicated those included in the "good savage" model, which Fortis had also invoked. Among them, honesty and a particularly strong sense of honor should be mentioned. The former trait went hand in hand with naivety, which was supposed to be an ally in the deception executed by Šćepan Mali. Zannowich pointed out that Montenegrins are characterized by a natural sincerity, which could be opposed with the corrupt morality of European high society:

> Slandering is not popular in Montenegro. This nation is too devoted to war, and too inactive to deal with the affairs of other people, there are also other reasons for that. Slander, so adored and acclaimed at one of the European courts, could not find a home in a country where the sounds of weapons were constantly heard, and where even the least important man would cut off the arms and legs of anyone who dared to slander him[6] (Zannovitch, 1784, p. 10).

---

[5] "La pureté des mœurs des Monténégrins surpasse de beaucoup celle qui existe aujourd'hui chez les nations civilisées. La jeunesse toute adonné aux armes, ne sent d'inclination pour le sexe que dans un âge mûr, au lieu que de nos jours l'homme & la femme se connaissent presque de l'enfance. Les filles élevées sous les yeux de leurs mères apprennent dans le silence, et loin de ces plaisirs séduisants qui amollissent aujourd'hui les jeunes cœurs, tout ce qui est nécessaire pour devenir des bonnes mères de famille. Rien n'est plus respecté chez ces nations que l'hospitalité et la vertu de femme."

[6] "La médisance qui trouve rarement des incrédules, & c'est un des privilèges du mal que l'on raconte, d'être cru sans preuves, & presque toujours sans examen, n'a point beaucoup de vogue dans le Monténégro. La nation est trop livrée aux dangers de la guerre, & trop peu oisive pour s'entretenir des affaires d'autrui à tort, & et à travers; d'autres raisons encore on pourrait alléguer : mais il nous suffira de dire que la médisance est si bien établie, & fêtée dans une certaine cour d'Europe, qu'elle ne se souci point de venir à s'y établir pas même pour quelques jours dans un pays qui retenait toujours le bruit des armes, & et où le plus petit individu lui couperait bras & et jambes si elle voudrait caqueter avec la même licence qu'elle fait ailleurs."

In fact this passage about "defamation" had a deep autobiographic aspect: Since Zannowich was an impostor himself, and pretended to be a prince of Albania, many treated him with suspicion and reserve. Zannowich decided to profit from the Western presumption on the natural righteousness and sincerity of peoples considered less civilized to pursue his own personal propaganda. However, it should be underlined that among all of the texts portraying the South Slavs that functioned among French readers of the second half of the 18th century, Zannowich's testimony was the most ambiguous, because he did not deny that Montenegrins are often deceitful, but pointed out that this applies mostly to their behavior toward foreigners, and is a consequence of their mistrust toward peoples other than Slavs (Zannovitch, 1784, p. 9).

Natural sincerity or, on the other hand, an inclination to lie, as an important element of the image of peoples considered less civilized, was an issue described not only in travelogues. The presumption that the Southern Slavs who inhabit provincial Dalmatia are simple, naturally good and sincere was also mentioned by a French official, André Abrial, who during the short period of Napoleonic rule in the Western Balkans (1806–1813) looked into the possibility of introducing the Napoleonic Civil Code in Dalmatia (Sajkowski, 2018, pp. 189–191). In his opinion, this task would be easy on the Dalmatian coast, because this part of the region, formerly belonging to Venice and to the Republic of Ragusa, was inhabited mainly by a literate population, which for centuries had been functioning according to different regulations based on Roman Law. The complex mosaic of different codes functioning in different cities was, of course, an obstacle to introducing one uniform Napoleonic Code, but Abrial thought this difficulty could be overcome. However, the inland part of the province was much more questionable when it came to introducing legal reforms. Although in theory the lands of Acquisto Nuovo, and Acquisto Nuovissimo, which the Venetians had gained from the Ottomans at the beginning of the 18th century, functioned according to the law of the Venetian Republic, the majority of the population was illiterate and lived according to their own common law which often contradicted the spirit of the Napoleonic Code (Sajkowski, 2018, pp. 189–190). The first problem stemmed from the illiteracy of the population, commonly referred to as the Morlachs, to whom the spoken

word held greater importance than written contracts. Therefore local priests, some of whom were literate and who were much respected by the local population, were meant to become officials of the state, as they could help with drawing up written contracts, e.g. marriage contracts. Apart from this main problem, others that were mentioned included the traditional patriarchal structure of Morlachian society, in which women were not treated as equal to men (which was a huge obstacle to introducing the institution of marriage as a civil contract). However, among all those inconveniences mentioned by Abrial, there was also one positive feature of the Morlachian character that would make legal reforms easier, i.e. natural sincerity. The French wanted to modify the whole structure of the judiciary left by Austria, which had ruled briefly over Dalmatia in 1797–1806 (Sajkowski, 2018, 192). The lowest level of the judiciary, i.e. justices of the peace, who were meant to resolve legal disputes of the first instance, was – according to Abrial – too severe. In his opinion, justices of the peace should be considered mild conciliators of disputes occurring among the locals, and not judges pronouncing harsh sentences. The conciliatory function of justices of the peace was possible to realize thanks to the Morlachs' natural sincerity. According to Abrial, the population of inland Dalmatia had no tendency for manipulation or litigiousness (Abrial, 1806).

On the other hand, when it comes to the practice of governance, the notion of sincerity has much wider implications than the above-mentioned aspect of judiciary reforms. It is also very important when it comes to the French governors' assessment of the loyalty of the Illyrian Provinces' populations. If Dalmatian Morlachs or the military settlers of the provinces of *Croatie Militaire* were naturally sincere, it meant that in case of disloyalty they would be unable to hide their hostility toward the French, or their commitment if their attitude toward their new rulers was positive. Hence reports in which French officers describing the disloyalty of the population mentioned examples of open hostility or collaboration with other political powers. For example, a report from Colonel Sorbier mentioned that almost every Dalmatian of the Orthodox confession has a portrait of the Russian tsar in his house (Sorbier, 1806). Other reports written by French administrators also mentioned examples of open hostility toward the French, which involved isolated acts of violent brutality or organized uprisings, partially provoked

by Russia (Sajkowski, 2018, pp. 116–135). On the other hand, the French were convinced that if the populations of the Illyrian Provinces were more favorably disposed toward French rule, they could become the most loyal soldiers of Napoleon because of the straightforward and devoted attitude that characterized all of their actions (Sajkowski, 2018, pp. 135–144).

Instead of ending this paper with a summary, it seems more apt to conclude with a quotation that not only sums up the experience of governance over the Dalmatian Slavs (or at least that part of the Slavic population which inhabited the provinces), but also sheds light on the previously mentioned testimonies on Slavic sincerity. Marshal Marmont, the first governor general of the Illyrian Provinces, wrote in his memoirs decades after that brief episode of the French presence in the Balkans:

> But this country, so sad and so poor, is inhabited by a nation so beautiful, so valiant and so susceptible to enthusiasm; so ignorant, simple, and confident, capable of devotion to its leaders; but, like all barbarians, it does not understand abstractions; to gain its favor, you have to awaken its senses and submit it to physical action[7] (Marmont, 1857, p. 27).

This testimony was published after Marmont's death in 1852, and it seems the 17th- and 18th-century presumption that there is some correlation between falseness and civilization still persisted. In this short quotation, confidence, devotion and sincerity are linked to an inability to understand abstractions and sophisticated truths of the civilized world. The "barbarians," as Marmont called the Slavs inhabiting provincial Dalmatia, were supposed to take all information literally, and anyone who wanted to gain their favor should also remember this. There is also a paradox which can be drawn from all of the above-mentioned testimonies: Uncivilized peoples' natural predisposition to sincerity and straightforwardness went hand in hand with their vulnerability to being fooled by those more civilized, while their natural truthfulness could easily be turned into credulity and superstition.

---

[7] "Mais ce pays, si triste et si pauvre, est habité par une population belle, valeureuse et susceptible d'enthousiasme; ignorante, simple, confiante, capable de dévouement pour ses chefs ; mais, comme tous les Barbares, elle ne comprend pas les abstractions; pour la remuer, il faut frapper ses sens et la soumettre à une action matériell."

References

Abrial, A. (1806). *Notice sur l'organisation, l'administration, la justice en Dalmatie, en aout 1806 - Sur l'application du Code Civil Napoleon A la Dalmatie, Rapport adressé de Zara le 28 septembre A S. M. J. Le Prince Eugêne Napoléon de France Vice Roi d'Italie par l'auditeur au conseil d'Etat de l'Empire Français Abrial*. Pierrefitte-sur-Seine: Archives Nationales AF IV 1713.

Bury, E. (1996). *Litterature et politesse, L'invention de l'honnête homme 1580–1750*. Paris: PUF.

Chastelet du, P. (1670). *La politique civile et militaire des Venitiens*. Cologne: Pierre Michel.

Duchet, M. (1971). *Anthropologie et histoire au siècle des Lumières*. Paris: Maspero.

Duval, P., Pyrard, F. (1679). *Voyage de Françoise Pyrard, de Laval: contenant sa navigation aux Indes orientales, Maldives, Moluques, & au Bresil: & les divers accidens qui luy sont arrivez en ce voyage pendant son séjour de dix ans dans ces pais: avec une description exacte des moevres, loix, façons de faire, police & gouvernement : du trafic & commerce qui s'y fait: des animaux, arbres, fruits, & autres singularitez qui s'y recontrent: divisé en trois parties (…) avec des Observations géographiques sur le présent voyage (…) par le Sieur du Val*. Paris: L. Billaine.

Ellingson, T. (2001). *The Myth of the Noble Savage*. Berkeley: University of California Press.

Fortis, A. (1778). *Travels into Dalmatia*. London: Robson.

Marmont, A. (1857). *Mémoires du Maréchal Marmont duc de Raguse de 1792 à 1841*, tome troisième. Paris: Perrotin.

Mortier, R. (2002). *Le «Prince d'Albanie». Un aventurier au siècle des Lumières*. Paris: Champion.

Motsch, A. (2001). *Lafitau et l'émergence du discours ethnographique*. Paris: Presses de l'Université de Paris-Sorbonne.

Petto, Ch. (2007). *When France Was King of Cartography: The Patronage and Production of Maps in Early Modern France*. Plymouth: Lexington.

Racault, J–M. (2003). *Nulle part et ses environs: voyage aux confins de l'utopie littéraire*. Paris: Presses de l'Université de Paris-Sorbonne.

Rapacka, J. (1997). *Leksykon tradycji chorwackich* [Lexicon of Croatian Traditions]. Warszawa: Slawistyczny Ośrodek Wydawniczy.

Sajkowski, W. (2013). *Obraz ludów bałkańskiego wybrzeża Adriatyku we Francji epoki Oświecenia* [Image of Peoples from the Balkan Adriatic Coast in France during the Enlightenment]. Poznań: Wydawnictwo Nauka i Innowacje.

Sajkowski, W. (2018). *French Image of the Peoples Inhabiting Illyrian Provinces*. Warszawa – Bellerive-sur-Allier: DiG-La Rama.

Sorbier (1806). *Dalmatie (reconnaissance du côtes de la), 4 juin 1806*. In: idem, *3 mémoires du colonel Sorbier sur les côtes et les îles de l'Istrie et de la Dalmatie, 4 et 6 juin 1806*. Vincennes: Service Historique de la Défense/ Divison Armée de Terre 1M 1590–1591.

Wolff, L. (2001). *Venice and the Slavs: The Discovery of Dalmatia in the Age of Enlightenment*. Stanford: Stanford University Press.

Watzlawick, H. (1999). *Bio-bibliographie de Stefano Zannowich*. Genève: Self-
-Published.

Zannovitch, S. (1784). *Notes historiques-politiques sur le Monténégro et les
Monténégrins, par un officier général qui a servi dans l'armée de l'empereur
Stiépan Mali, l'année 1778 et jusqu'à sa mort arrivée le 15 août 1779 dans
la province de Czernizza*. In: idem, *Stiepan-Mali, c'est à dire Etienne Petit,
ou Stefano-Piccolo.,le Pseudo Pierre III. Empereur de Russie, qui parut dans
le grand-Duché de Monténégro, situé entre la mer Egée, l'Albanie Turque &
le golfe Adriatique, en 1767, 1768 & 1769*. Mangalor, forteresse du Nabab
Hyder–Haly, sur les côtes du Malabar (Vienne?).

Adam Grzeliński
ORCID 0000-0002-4007-6507
Institute of Philosophy, Nicolaus Copernicus University in Toruń

The Validity of Aesthetic Judgments: George Santayana's Polemics with Tradition

Abstract

The paper discusses issues of the validity of aesthetic judgements in the concept of George Santayana, the author of *The Sense of Beauty* and *Reason in Art*. Although Santayana places aesthetic experience as well as experience as such at the center of his analyses, and explicitly opposes the view that aesthetic judgements are supra-individual, I endeavor to demonstrate that his concept is not necessarily contrary to the achievements of 18th-century aesthetics proclaiming the validity of judgements of taste (Shaftesbury, D. Hume, E. Burke or I. Kant). It is true that Santayana's multifaceted understanding of beauty reveals the limitations of Kant's aesthetics, but nevertheless, it does not contradict its achievements. The differentiation he made between the beauty of matter, form, and expression suggests a dual meaning of the validity of aesthetic judgements, on the one hand pointing to their universality while on the other referring to the ideals of reason whose sensual representation is available in works of art.

Keywords: aesthetic judgement, aesthetic experience, Santayana, Kant

Classical philosophical aesthetics, flourishing in the 18th century which is sometimes referred to as the "age of taste" (Pazura, 1981), developed several solutions to the problem of justifying the validity of aesthetic judgments. The very problem of the validity of these judgments was connected with the emergence of a new understanding of human subjectivity, in which the individualism of human beings was emphasized more clearly than before: Although the human was traditionally referred to as *animal rationale*, reasonableness distributed among individuals proved to be flawed, and only the collective effort of particular researchers enabled universally accepted knowledge to be attained – but in

order to reconstruct this process of rising toward reasonableness, it was necessary to use individual experience as a starting point. Also human affectivity, analyzed with ever growing accuracy, showed human persons whose passions placed them not only in relation to the world of nature, but also to other people. Finally, this discovery of the individual was crowned with a practical culmination in the form of a social and political doctrine with slogans referring not only to the natural law binding for all people, but also to the social contract posited by John Locke and Jean-Jacques Rousseau.

In aesthetics, such a perception of the individuality of humans led 18th-century thinkers to enquire about the principle underlying aesthetic judgements: They stem from personal experience, but at the same time they are not limited to it and have validity, something that at the end of the century Kant would call "the universal validity claim" (Kant, 1914, p. 61). Thus, although it was noted that the particular pleasure inherent in the perception of beauty and associated with the aesthetic disinterestedness discovered at the beginning of the 18th century is subjective, various justifications were sought to distinguish said subjectivity from individuality and to guarantee that aesthetic judgements built on this pleasure could be valid. The distinction between individual and objective predilection or, at least, the intersubjective status of a judgement, was to serve as a condition for any discussion about beauty and other values of sensual experience such as sublimity or picturesqueness. In the British aesthetics of the 18th century, the search for justification of the validity of judgements went in several major directions determined by the achievements of three philosophers and aestheticians: Anthony Ashley Cooper Shaftesbury, Edmund Burke, and David Hume. Although it is easy to point to the legitimacy of the problem they were trying to quantify, the solutions they proposed were only partially satisfactory. Although the Platonizing concept of Shaftesbury contained the discovery of aesthetic disinterestedness, beauty was understood in it in an extremely broad sense (at least as a combination of beauty and sublimity, but also of inner beauty, i.e. virtue) – ultimately all phenomena that were capable of arousing sincere admiration were described by Shaftesbury as beautiful, as they allow us to understand the purposive organization of nature and contribute to the formation of an appropriate inner human attitude.

On the other hand, although Burke's search for natural (physical, biological, physiological and psychological) determinants of the objectivity of judgements of taste allowed two fundamental aesthetic values to be distinguished already at the start, namely beauty and sublimity, and helped distinguish the purposiveness of beauty from the purposiveness of life, this nevertheless led to the arbitrariness of solutions (Grzeliński, 2001, p. 114). The description of the physiological mechanism by which objects with certain properties influence the human perception apparatus did not leave any doubt – judgements of beauty and sublimity proved to be equally indispensable as judgments of other properties of objects. Burke's attempt to explain the discrepancies of judgements with the idea that habit changes the natural response, revealed the precarious nature of the entire structure. Natural taste, common to all, turned out to be only a postulate, and its verdicts were uncertain. Finally, the third solution, proposed by Hume, described the forming of an opinion on works of art and how such an opinion functions in society rather than its actual subject (Hume, 1955). In fact, Hume's concept describes how artistic values are defined through the formation of opinions and habits, but it does not answer the question of what beauty actually is, and how artistic beauty differs from natural beauty.

The 18th-century discussions were summarized, so to speak, in Immanuel Kant's concept presented in his *Critique of Judgement* (1790). The answer it provided was based on previous discoveries, yet surpassed them by indicating a completely different way of establishing the conditions for the validity of aesthetic judgements. In the case of beauty, this condition was spatial order or, in the case of music, temporal order. The procedure proposed by Kant was to reduce the phenomenon of beauty to its simplest form in which one can answer the question about the basis for the common inclination of all people. In his *Analytic of the Beautiful*, Kant pointed to four "moments" of judgement of taste which are derivatives of particular groups of categories that allow one to formulate any judgement: disinterestedness (quality category), universality (quantity category), purposiveness without purpose (relationship) and necessity (modality). Thus, an object was to be considered only as a phenomenon, i.e. in isolation from its existence (which could be the cause of either direct sensual stimulation and the related pleasure, or satisfaction resulting from the possibility of making use of the object

for practical purposes). Ultimately, beauty was reduced only to the form of a phenomenon – to the order found in the pure intuition of time and space imposed on phenomena, disregarding the matter of the phenomenon – the charm connected with the quality of phenomena, with color or tonality of sound. Because all people are gifted with the capability to perceive phenomena in time and space as well as the rudimentary ability to find order in such perceptions, one should expect that in the case of thus understood beauty, judgements of aesthetic taste may claim the right to universality. Ultimately, therefore, while people may differ in their judgments when guided by personal inclinations, aesthetic judgements may be accorded the value of validity.

Santayana's objection

From this perspective, the starting point for the concept presented by George Santayana (1863–1952) in his two works from the turn of the 19th and 20th centuries, *The Sense of Beauty* (1896) and *Reason in Art* (1905), is perhaps a contradiction to all of the above findings. In the introductory paragraphs of the former work, the philosopher insists on abandoning the concept of disinterestedness which constituted a foundation for all the aesthetic considerations of 18th-century aesthetics, as well as the aspirations to establish any kind of standard of taste desired by Hume, or the universality of the principles of beauty.

> It is unmeaning to say that what is beautiful to one man ought to be beautiful to another. If their senses are the same, their associations and dispositions similar, then the same thing will certainly be beautiful to both. If their natures are different, the form which to one will be entrancing will be to another even invisible, because his classifications and discriminations in perception will be different, and he may see a hideous detached fragment or a shapeless aggregate of things, in what to another is a perfect whole – so entirely are the unities of objects unities of function and use. It is absurd to say that what is invisible to a given being *ought* to seem beautiful to him. Evidently this obligation of recognizing the same qualities is conditioned by the possession of the same faculties. But no two men have exactly the same faculties, nor can things have for any two exactly the same values (Santayana, 1896, p. 27).

Santayana's negation of the universality of judgements of taste goes hand in hand with his opposition to the possibility of aesthetic

disinterestedness. In his interpretation, disinterestedness would imply a specific mode of experience. Much of our day-to-day relation to the world around us is "interested" in character (from the Latin *inter esse*), i.e. related to the existence of objects, their usefulness, the possibility of manipulating them or avoiding them in case of danger; contrary to this, a disinterested approach to the world would mean abstraction from the real existence of things that surround us, a sort of gliding of the soul delighted by beauty over material reality. Such an understanding of disinterestedness seems, to the American philosopher, to be at most the result of a momentary reflection, and ultimately an illusion, since it is impossible to break the bond with the body situating us amidst other bodies. If disinterestedness is to be understood as a reference to sensual pleasure as such, without looking at the consequences, then in fact every pleasure felt would be selfless in character. Santayana's argument finds its culmination in the statement that disinterestedness cannot constitute the *differentia specifica* of an aesthetic experience.

There seem to exist at least two reasons for such an interpretation. One of them is Santayana's shift toward experience, that "invigorating sense that here, at last, we are getting back to experience, down to earth," as Monroe Beardsley, another American aesthetician, wrote in his commentary on his concept (Beardsley, 1975, p. 329). Indeed, Santayana is more interested in the experience itself, in the "sense" of beauty as indicated by the title of his work, and not in aesthetic theory; as he puts it, "to feel beauty is a better thing than to understand how we come to feel it" (Santayana, 1896, pp. 8–9). The second reason is an attempt to overcome the limitations of traditional aesthetics, which is expressed in Kant's thought. Above all, he notes that Kant's effort to define the conditions for aesthetic judgements (i.e. statements that would merit attribution of the value of validity) leads to a significant impoverishment of aesthetic experience. This is because it is limited only to the sphere of formal beauty – the pleasure derived from experiencing spatial or temporal order. Indeed, this experience refers to a specifically aesthetic sphere, as it cannot be explained through reference to another type of experience, yet it is not the only aspect of aesthetic experience.

> Where there is a sensuous delight, like that of color, and the impression of the object is in its elements agreeable, we have to look no farther

for an explanation of the charm we feel. Where there is expression, and an object indifferent to the senses is associated with other ideas which are interesting, the problem, although complex and varied, is in principle comparatively plain. But there is an intermediate effect which is more mysterious, and more specifically an effect of beauty. It is found where sensible elements, by themselves indifferent, are so united as to please in combination. There is something unexpected in this phenomenon, so much so that those who cannot conceive its explanation often reassure themselves by denying its existence (Santayana, 1896, p. 53).

Formal beauty, the explanation of which is a unique task of aesthetics, is an intermediate phenomenon between material beauty and the beauty of expression. The former is related to the satisfaction of simple sensual qualities – color, sound, or touch. Satisfaction in them is the simplest manifestation of aesthetic experience, not requiring any training. Such an experience is contrary to the postulates of earlier aesthetics: It is individual, and often, as in the case of tactile properties, it is a result of the human body's interaction with the environment, and therefore it is not disinterested in the sense indicated above. Acknowledgment that this kind of experience is aesthetic in character goes hand in hand with negating the selfless character of aesthetic experience and the claim of validity of aesthetic judgements. In other words, according to Santayana, the relegation of the material aspect of this experience from the sphere of beauty was too high a price that Kant had to pay for determining the conditions of the universality (i.e. validity) of aesthetic judgements.

At the opposite pole of beauty, so to speak, we find the beauty of expression. While the beauty of matter is an expression of a simple, spontaneous inclination characteristic of both children and primitive people, who "delight in bright and variegated colors …, appreciate the neatness of muslin curtains, shining varnish, and polished pots" (ibidem, p. 50), the beauty of expression is subject to formation. It concerns a specific effect when a phenomenon appropriately expresses the content associated with it, while the source of satisfaction is not the pleasure awakened by the phenomenon itself (we would then be dealing with the beauty of matter or form) nor the nice content itself. It concerns the adequacy of this association. For although we are dealing with expression in various forms of communication, if it is to be an element of beauty, it must meet a certain condition:

I may see the relations of an object, I may understand it perfectly, and may nevertheless regard it with entire indifference. If the pleasure fails, the very substance and protoplasm of beauty is wanting. Nor, as we have seen, is even the pleasure enough; for I may receive a letter full of the most joyous news, but neither the paper, nor the writing, nor the style, need seem beautiful to me. Not until I confound the impressions, and suffuse the symbols themselves with the emotions they arouse, and find joy and sweetness in the very words I hear, will the expressiveness constitute a beauty; as when they sing, *Gloria in excelsis Deo* (ibidem, p. 122).

But is Santayana's position truly individualistic, and is the entire aesthetic experience relative to the individual and untransferable? If this were the case, we would be faced with a complete rejection of the achievements of earlier aesthetics, of the above-mentioned attempts to establish conditions for the validity of aesthetic judgements. I believe that while Santayana expands the plane of experience of aesthetics and makes the process of experiencing rather than the act of judging and the character of judgements the subject of his analyses, these two approaches – the pursuit to establish the conditions for the validity of judgements of taste (as sought by such philosophers as Shaftesbury, Hume, Burke or Kant) and the individualism of experience, on which Santayana insisted – can be reconciled to some extent.

First of all, let us note that Santayana's starting point is naturalistic and empirical. Similarly to William James, whose psychology lectures he attended, he believes that experience in its original state is an undifferentiated total of available data, from which certain elements become objectivized only through the activity of the subject: even though the natural boundary between the subjective and the objective sphere of experience is fluid. The fact that historical philosophical concepts defined this boundary in different ways stems from the fact that this action, and thus the attention paid to it, causes, as James wrote, "substantial" and "transitive" components to be distinguished in experience (James, 1890, p. 243 ff.). The said attention led James to postulate in his later work the existence of a completely undifferentiated, "pure experience," not yet completely polarized into the subjective and objective sphere (James, 1912, p. 39 ff.). Although some part of experience is objectified in the process of life, there is a kind of experience in which the process is not complete, leaving a special kind of "intermediate sphere":

There is a peculiar class of experiences to which, whether we take them as subjective or as objective, we assign their several natures as attributes, because in both contexts they affect their associates actively, though in neither quite as 'strongly' or as sharply as things affect one another by their physical energies. I refer here to appreciations, which form an ambiguous sphere of being, belonging with emotion on the one hand, and having objective 'value' on the other, yet seeming not quite inner nor quite outer, as if a diremption had begun but had not made itself complete (ibidem, p. 34).

Classifications depend on our temporary purposes. For certain purposes it is convenient to take things in one set of relations, for other purposes in another set. In the two cases their contexts are apt to be different. In the case of our affectional experiences we have no permanent and steadfast purpose that obliges us to be consistent, so we find it easy to let them float ambiguously, sometimes classing them with our feelings, sometimes with more physical realities, according to caprice or to the convenience of the moment (ibidem, p. 141–142).

The "masterly way" in which, according to James, Santayana explores the subject leads him to define beauty as an objectified pleasure: "beauty is pleasure regarded as the quality of a thing" (Santayana, 1896, p. 31). But since it is only "convenience and economy of thought alone [that] determine what combination of our sensations we shall continue to objectify and treat as the cause of the rest" (ibidem, p. 30), aesthetic experience proves to be a sphere in which such a classification has not yet taken place, as it is not connected with any practical purpose. Thus, Santayana reverses the former approach to describing aesthetic disinterestedness in a manner characteristic of James' radical empiricism. It does not involve suspending references to existing objects in order to be able to admire the phenomenon itself; on the contrary: in this case, experience is not yet divided into the subjective and the objective sphere, while aesthetic experience in a way retains the remains of its original, undifferentiated state. It seems, therefore, that Santayana is actually describing the same mode of experience that we found earlier in Shaftesbury (Shaftesbury, 2001, p. 236; Stolnitz, 1961, p. 131 ff.) and which later became the basis for determining the specificity of aesthetic judgements in Kant (Kant, 1914, p. 48).

The same is true of the second objection raised by Santayana against the aesthetic tradition. Emphasizing the primacy of experience in relation to theory indeed allows us to draw attention

to the difficulties associated with the expectation that the same objects should be the source of aesthetic pleasure for everyone. Even if previous attempts to determine the conditions for the objectivity of aesthetic judgements were the result of an erroneous approach to aesthetics, it was a mistake that was too widespread, too persistent and too long-lasting to be rejected without further reflection. The project of determining a standard of taste or seeking grounds for the validity of aesthetic judgements seems to play an important role also in Santayana's concept. Since all aesthetic judgements originate in individual experience, every judgement of taste will be an expression of individual inclination. In that case, the claim of validity of an aesthetic judgement would imply that it is possible to find a correspondence between a judgement expressing an attribution of an aesthetic value to a certain object and a possible judgement expressed by someone else. The compatibility of such judgements does not imply that the object evokes exactly the same pleasure in two individuals – as this cannot be verified – but that each of them acknowledges that the object is beautiful and that it can potentially be regarded as beautiful by others as well. Thus, although the judgement on the beauty of an object may initially be unclear, because it may not find confirmation in someone's individual experience, it makes it possible to enhance that experience and develop one's own taste, even if the experience always remains individual.

Back to traditional aesthetics

Let us note, however, that Kant's claim to the universality and validity of aesthetic judgements has a different meaning in relation to the three types of beauty indicated by Santayana: material, formal, and the beauty of expression. With regard to the first of these, it is indeed difficult to speak of specifying the conditions for the possibility of issuing aesthetic judgements: The predilection for what Kant calls the matter of the phenomenon, i.e. the aspect of experience that is qualitative and cannot be reduced to spatial and temporal order, is individual in character. This fact is confirmed by the multitude of physiological, psychological and social factors indicated by Santayana which could determine whether someone finds pleasure in one sensual experience or another (Santayana, 1896, p. 27). While we may point to some rules constituting generalizations of experience,

they lack the character of an actual determination. Slightly different is the case with the beauty of a form. As indicated by Kant, the claim of validity of judgments of taste results from two circumstances: the universality of sensual perception in space and time, and the possibility given to everyone to discover the order within them. As the matter concerns spatial and temporal relations, it is possible to determine whether a given form is refined to a greater or lesser extent. Such a "standard of taste" must, according to Santayana, still refer to experience and ultimately find its confirmation in the fact that it conveys aesthetic pleasure to someone. In this case, as we have just noticed, one can point to the propaedeutic role of aesthetic judgements that allow for the development and formation of individual aesthetic taste.

The situation becomes more complicated in the case of the beauty of expression. When describing aesthetic phenomena, Kant emphasizes the difference between beauty and the beauty of a symbol or expression: "Beauty is the form of the purposiveness of an object," as we read in *The Critique of Judgment*, "so far as this is perceived in it without any representation of a purpose" (Kant, 1914, p. 90). Such a presentation of a purpose may concern either an objective internal purpose (in the case of living beings) or an objective external purpose (in the case of human products such as tools). The separation of the two orders introduced by Kant allows us to distinguish unique aesthetic pleasure from satisfaction felt in relation to the appropriateness and handiness of a tool. Although both types of purpose – subjective related to the appropriateness of the form of a phenomenon in relation to human cognitive powers (i.e. the possibility of discovering order among phenomena) and objective (related to the possibility of defining a specific purpose) – may accompany each other, it is only in the case of "purposiveness without purpose" that the experience of free beauty, i.e. beauty not related to any concept, is possible. "The first is called the (self-subsistent) beauty of this or that thing; the second, as dependent upon a concept (conditioned beauty), is ascribed to Objects which come under the concept of a particular purpose" (ibidem, p. 81). Free beauty is only abstract; conditioned beauty is a combination of beauty (always abstract) and the appropriateness of the object in relation to its concept (in this sense one can talk about the beauty of a building: its form is structured, while at the same time it must meet the requirements of being a building). The combination of concept

and beauty makes it possible to define an object as beautiful (e.g. a beautiful building) only when it corresponds to a given concept and additionally represents a certain organized whole.

However, a unique kind of conditioned beauty is found in artistic creativity. As Kant puts it, it is a beautiful representation of a thing (ibidem, p. 193), which seems to correspond to what Santayana calls the beauty of expression. Kant, however, does not discuss it in *Analytic of Beauty* but in the paragraphs of *The Critique of Judgement* that relate to fine arts. This is probably due to the fact that while in the case of free (formal) beauty one can point to general human conditions for making judgements about it, in the case of artistic beauty this is much more difficult, since it involves various cognitive factors such as the education of the audience, their upbringing, etc. In this case the question of the validity of aesthetic judgements gains different meaning.

It should be noted that when analyzing aesthetic phenomena, Kant distinguishes between two types of concepts, which enables the problem of what he calls the antinomy of taste to be solved. This issue is closely related to the problem of validity of aesthetic judgements – in relation to both Kant's and Santayana's concepts. Namely, Kant refers to everyday observation, on the basis of which perhaps aesthetic judgements cannot be proved, but may be the subject of disputes:

> There emerges therefore in respect of the principle [*Prinzip*] of taste the following antinomy:
> 1. Thesis. The judgement of taste is not based upon concepts; for otherwise it would admit of controversy (would be determinable by proofs)
> 2. Antithesis. The judgement of taste is based on concepts; for otherwise, despite its diversity, we could not quarrel about it (we could not claim for our judgement the necessary assent of others) (ibidem, p. 231).

As we know, the resolution of this antinomy does not boil down to stating that one of the theses presented is true, but to demonstrating that the meaning of the term "concept" is understood differently in each of them. If a concept was purely intellectual in character, and thus could be defined, we would be dealing with a cognitive judgement and could decide whether a given object is a representation of a given concept (based on this principle it is possible to decide whether a given object is e.g. a building).

The judgement of taste is not based on such concepts. However, the antithesis is about a slightly different kind of concept: "In the Thesis we mean," says Kant, "that the judgement of taste is not based upon determinate [*bestimmten*] concepts; and in the Antithesis that the judgement of taste is based upon a concept, but an *indeterminate* one (viz. of the supersensible substrate of phenomena). Between these two there is no contradiction" (ibidem, p. 234). Mirosław Żelazny, the Polish translator of the new edition of *The Critique of Judgement*, adds this explanation: "The supersensible substrate of phenomena here means a pure representation of perfect beauty as a model for assessing the perfection of a sensual object considered beautiful in a way that is closer to or further from that perfection" (Kant, 2014, p. 220). In the case of free beauty this would mean comparing the form of a phenomenon with a pure representation, e.g. an organized geometric form whose empirically given shape would be only an approximation. Assessing beauty would mean comparing this empirical form with a pure representation, for example with some regular geometric shape.

A slightly different concept of beauty is applied in the context of artistic work. The concept called the supersensible substrate of phenomena is then synonymous with a different concept: an aesthetical idea, "an intuition (of the imagination) for which an adequate concept can never be found" (ibidem, p. 236). "I believe," Kant adds, "we might call the aesthetical idea an *inexponible* representation of the imagination" (ibidem). This 'thought-provoking' aesthetical idea, which points to certain content but at the same time cannot be included in a conceptual definition, occurs in fine arts, which are the domain of human genius.

This particular field, the domain of art revealing the ideals of human reason, allows us to articulate what Santayana calls the beauty of expression. To reiterate: We deal with it when there occurs an adequate representation of certain content in sensual form. The form itself does not necessarily have to be beautiful, just as the content it refers to does not have to evoke pleasant associations. The essence of the beauty of expression is only the appropriateness of the way in which the content is presented, in such a way that it is actually revealed. The beauty of expression is at the same time a symbolic beauty. Santayana's naturalism goes hand in hand with his specific idealism: Earlier positions (for example, Shaftesbury's concept in the field of aesthetics)

were erroneous when considering reasonableness as the basis of existence, whereas in fact it constitutes its purpose. The purposiveness that governs nature, which the author of *The Moralists* sought in the universal dependency of things and in the admiration of nature, can at best be the purposiveness of human reason. According to Santayana, aesthetic perception of nature, admiration of the purposiveness permeating all phenomena, is the result of imposing reasonable intentionality on the very perception of phenomena. Art, among other things, is the proper and primary area of revealing the ideals of reason; although its origins are completely natural and strongly connected with the biological aspect of human life, its crowning achievement is the sensual representation of the ideals of reason (Santayana, 1905, pp. 34–35).

> This fluidity of the mind would make reflection impossible, did we not fix in words and other symbols certain abstract contents; we thus become capable of recognizing in one perception the repetition of another, and of recognizing in certain recurrences of impressions a persistent object (Santayana, 1896, p. 119).

With regard to beauty of expression understood in this way, validity should be considered differently than in the case of the universality of inclination indicated earlier. The measure of validity of an aesthetic judgement is no longer the possibility of its confirmation by others, but the judgement's pronouncement of the adequacy of the representation of an aesthetic idea in a work of art. At this point, Santayana's comments, although verbally opposed to the traditional understanding of aesthetics with its concepts of disinterestedness or universality of judgement, surprisingly seem to support Kant's remarks on beauty and artistic work. While in the case of free beauty (beauty of form), Kant points to the possibility of justifying the validity of judgements of taste, in relation to aesthetical ideas created and developed by art, he speaks only of the possibility of their understanding by others. By the same token, over a hundred years later Santayana seems to point to the impossibility of indicating the reasons determining a judgement, and thus the impossibility of imposing its understanding on everyone. By analogy, Santayana does not prejudge that, by definition, a judgement referring to artistic beauty must remain individual, but emphasizes the possibility of developing aesthetic sensitivity. However, this is not only a matter of the capacity for increasingly accurate recognition of order in the

apparent chaos of sensual experiences (as in formal beauty), but also an increasingly better understanding of the ideals of human reason, for which art is a natural way of symbolic representation. Just as in the beauty of form the judgments of others allowed us to broaden our own experience, the beauty of expression allows us to enhance not only our experience, but also our understanding of the human world co-created and co-inhabited with others.

References

Beardsley, M. C. (1975). *Aesthetics from Classical Greece to the Present*. Tuscaloosa: University Alabama Press.

Burke, E. (1823). *A Philosophical Inquiry into the Origin of Our Ideas of the Sublime and Beautiful*. London: Thomas M'Lean.

Danto, A. (1988). An Introduction. In: G. Santayana, *The Sense of Beauty. Being the Outlines of Aesthetic Theory*, W. G. Holzberger, H. J. Saatkamp Jr. (Eds.) (pp. 3–11). Cambridge, MA: MIT Press.

Grzeliński, A. (2001). *Angielski spór o istotę piękna. Koncepcje estetyczne Shaftesbury'ego i Burke'a* [The English Dispute over the Essence of Beauty. The Aesthetic Concepts of Shaftesbury and Burke]. Toruń: Wydawnictwo Adam Marszałek.

Grzeliński, A. (2017). Naturalne źródła twórczości artystycznej w koncepcji estetycznej George'a Santayany [Natural sources of artistic creativity in the aesthetic concept of George Santayana]. In: A. Nowak (Ed.), *Interpretacje dzieła muzycznego* [Interpretations of a Work of Music] (pp. 23–35). Bydgoszcz: Wydawnictwo Akademii Muzycznej.

Hume, D. (1987). Of the standard of taste. In: D. Hume, *Essays Moral, Political, and Literary*, E. F. Miller (Ed.) (pp. 226–249). Indianapolis: Liberty Fund.

James, W. (1890). *Psychology*. New York: Holt.

James, W. (1912). *Essays in Radical Empiricism*. Mineola: Dover Publications.

Kant, I. (1914). *The Critique of Judgement*, transl. J. H. Bernard. London: Macmillan.

Kant, I. (2014). *Krytyka władzy sądzenia* (Polish edition of *Kritik der Urteilskraft*, transl. M. Żelazny). Toruń: Wydawnictwo Naukowe UMK.

Pazura, S. (1981). *De gustibus. Z dziejów pojęcia smaku estetycznego* [De gustibus. From the History of the Notion of Aesthetic Taste]. Warszawa: Państwowe Wydawnictwo Naukowe.

Santayana, G. (1896). *The Sense of Beauty*. New York: Scribner.

Santayana, G. (1905). *The Life of Reason; or, The Phases of Human Progress. Vol. 4: The Life of Reason*. New York: C. Scribner's Sons.

Shaftesbury, A. A. C. (2001). The Moralists. In: A. A. C. Shaftesbury, *Characteristics of Men, Manners, Opinions, Times*, D. den Uyl (Ed.) (pp. 101–247). Indianapolis: Liberty Fund.

Stolnitz, J. (1961). On the origins of aesthetic disinterestedness. *The Journal of Aesthetics and Art Criticism* XX, No. 2, 131–143.

Anna Żymełka-Pietrzak
ORCID 0000-0003-1474-2358
»Artes Liberales« Academy

Not Naked but Wearing "Dress upon Dress": Johann Georg Hamann on Truth

Abstract

The chapter presents a detailed analysis and interpretation of Johann Georg Hamann's metaphor of truth clothed in dress upon dress, which was introduced in Hamann's letter to Kant of July 27, 1759. Truth is displayed as a woman in many layers of dresses, but removing the clothing reveals a fearful ghost. This metaphor is in the chapter regarded as emblematic for Hamann's reflection on truth and has been interpreted in the light of the author's theologico-ethical model in which God's truth becomes accessible to humans in communicational acts of condescension. The incarnation of Jesus Christ is the most important act of this kind, and comprises a theological matrix for human authorship. The incarnation of truth is an imitation of God's kenotical act and is understood as service to the truth.

Keywords: Johann Georg Hamann, truth, incarnation, condescension, authorship

1. Hamann's metacritique

In his thinking about truth, Johann Georg Hamann, a German author living in Königsberg, Prussia, in the late Enlightenment, was neither homogeneous nor unambiguous (Stünkel, 2007, p. 156). It was his deliberate strategy to give priority to metaphors and poetical language, since he claimed that the theoretical language of concepts, making use of abstractions, is so significantly detached from the historical concrete that it becomes an "empty wineskin" (N, III, 285, 36; transl. G. Griffith-Dickson, 1995, p. 521)[1] which

---

[1] Letters edited by W. Ziesemer and A. Henkel (1955–1979) as well as the critical edition of Hamann's *Collected Works* edited by J. Nadler (1949–1957) are quoted according to the format established in the scientific literature on the

can be arbitrarily reshaped and inflated by philosophers according to their own interests and liking. In an often cited excerpt from a letter to Jacobi, this Prussian thinker distances himself from systematic philosophy: "System is already in itself a hindrance to truth" (ZH VI, 276, 15f, transl. W. M. Alexander, 1966, p. 49). Hamann objects to the inclination for system building, and his skepticism is aimed at the principle of generality, the principle of non-contradiction (see: Moustakas, 2003, pp. 176–178)[2] and the claim for totality. They all seem to remain in conflict with the concrete of historical, sensorial, individual existence.

Hamann charges systematic philosophy with purism: striving to purify thinking of tradition, faith, cultural transmission, everyday experience, and, finally – of language, along with its contingent, historical characteristics. To prove the illegitimacy of such purification, Hamann wrote *Metacritique of the Purism of Reason* (N, III, 281–289; transl. G. Griffith-Dickson, 1995, p. 517–525). Griffith-Dickson (1995) takes this term, *metacritique*, in the broad sense, as the main characteristic of the whole of Hamann's way of thinking and writing. Hamann's writings are metacritical through and through, they are oriented toward language as the medium of any philosophical effort. To the Magus of the North, language is the medium of thought – and its various "contaminations" are indispensable and essential for creative philosophizing and for expressing the truth. Language is the space of imagination, a deposit of images and vivid metaphors, and as such it comprises the organon of thinking (Regier, 2016, pp. 171–173). Metacritical philosophizing, consciously respecting the nature of language, fulfills itself in drawing upon this deposit as well as in investigating hidden metaphors underlying opponents' argumentation, and disclosing their unobvious consequences[3].

----

topic, i.e. ZH VI, 276, 15 (for correspondence) or N, I , 45, 19 (for works), where after the letter symbol comes the volume number, page number and line. After that comes the quotation of the English translation.

[2] In *Metacritique of the Purism of Reason* he writes that metaphysics "misuses all the word-signs and figures of speech of our empirical knowledge as lukewarm hieroglyphs and types of ideal relations, and with this learned misdemeanour re-works the **straightforwardness** of language into such a senseless, ruttish, unsteady, imprecise Something = x, that nothing remains but a windy sighing, a magic shadow-play" (N, III, 285, 28–34, Hamann, 1995, p. 521, transl. G. Griffith-Dickson).

[3] For instance, in his review of *Critique of Pure Reason* Hamann takes into consideration Kant's metaphor of the tribunal of reason mentioned in the preface

2. Blumenberg's metaphorology
as a methodological approach

Hans Blumenberg's (2010) metaphorological approach seems to be useful for investigating Hamann's metaphors in the course of intellectual history. Metaphorology is conceived as an auxiliary discipline for the history of concepts. It aims to analyze absolute metaphors that cannot be reduced to concepts, to a logical form, and at the same time comprise the very *sub-structure of thinking* and determine the field of philosophical terminology (Blumenberg, 2010). According to Blumenberg as well as Hamann, a complete translation of metaphors into theoretical language is impossible and also undesirable. Metaphors resist translation into concepts.

My aim is, however, to elucidate Hamann's metaphors in the light of some of his other writings, as this could highlight their importance for his thinking about truth against the background of his theologico-philosophical worldview. I am going to consider some of the philosophical and theological consequences of employing a certain metaphor. Metaphorology is a promising way of doing this, since it is able to show one of the most interesting moments of Hamann's writing, namely the tension between logico-theoretical and poetic language.

Hamann thinks about truth through images: Truth is a seed which can be sown, it is the sun or a little light, or even a metal ("Truths are metals which develop under the earth"; ZH, I, 304, 19–20; transl. W.M. Alexander, 1966, p. 173). Often truth is depicted as a woman, daughter, beloved, or a wife to whom someone gives a bill of divorce. Those images are rarely translated into concepts by Hamann himself, although they entail complex structures of thinking and can be developed as a result of hermeneutical effort.

I am going to investigate the metaphor of truth in which truth is presented as a woman wearing many dresses at once, in many layers. I suggest understanding this metaphor as an expression of such a concept of truth in which truth conceived theologically is available to human beings in its historical and linguistic manifestations. Hamann's perspective is not epistemological but

---

to the *Critique*. In a letter to Ch. J. Kraus from Dec. 18, 1784 (Hamann, 1996), he analyzes Kant's metaphor of immaturity and guardianship included in the essay *An Answer to the Question: What Is Enlightenment?*

ethical; in it, the proper attitude toward truth is an attitude of service and personal engagement.

3. Biographical background (*Sitz im Leben*)

The metaphor which captured my interest, presenting the truth as clothed in many garments, is found in the end part of Hamman's letter to Immanuel Kant dated July 27, 1759 (ZH I, 373–381, Hamann, 1967). Hamann wrote this letter just two days after his first meeting with Kant, which had been arranged by their common friend, Johann Christoph Berens. Two years earlier, hired by the Berens family at the time, Hamann set off for London with a secret trade mission, perhaps not without a political aspect – as the Seven Years' War had already been going on for a year. The mission, which itself remains fairly obscure, ended in crushing failure and Hamann, left without friends and allies, with no means of providing for himself, went through a religious awakening whose causes he himself associated with careful reading of the Bible through the lens of his personal experience. From then onward he began to interpret his life in the light of biblical stories and considered his authorship [*Authorschaft*] as a vocation leading him to the continuation of divine revelation through literary activity. In 1758 Hamann returned, first to Riga, then to Königsberg, where he remained idle at his family home.

In the aforementioned letter Hamman refers critically to the efforts of Kant and Berens, who tried, after coming to the meeting with Hamann from Riga to Königsberg, to dissuade him from his newly rediscovered Lutheran spirituality and convert him to the much safer, rational, enlightened deism. As a means of accomplishing this conversion, Kant proposed that Hamann translate excerpts from the *Encyclopédie* about beauty, art, and serfdom from the French. Hamann rejected the proposal by letter, expressing great praise of faith which, on the basis of an idiosyncratic interpretation of Hume's philosophy, he considered necessary to conduct any act in the world.

The statement about truth appears in the context of a critical assessment of Enlightenment philosophy. Hamann, as he often indicated, perceived Enlightenment philosophy as being guilty of pride and idolatrous – deifying reason instead of worshipping the true God.

4. Analysis and interpretation of the metaphor

In a broad context, the metaphor in question reads as follows:

> You know, most honored tutor, that genii have wings and that they sound just like the applause of the multitude. If one is permitted to mock God with grace and strength, why shouldn't one be able to amuse oneself with idols? Mother Lyse sings: Make mockery of idols false. A philosopher however looks at poets, lovers, and visionaries as a man looks at a monkey, with pleasure and pity. As soon as men can understand one another, they can work. He who confused the languages—who punished the exemplars of pride out of love and also for the sake of political ends, for the good of the populace as a friend of humanity—joined them together again on the day that they slandered men with tongues of fire, as if intoxicated by sweet wine. The truth did not want highway robbers to get too close to her; she wore dress upon dress, so that they had misgivings about ever finding her body. How terrified they were when they had their wish and saw the truth, the terrible ghost, before them (ZH, I, 380,34–381,11; transl. A. Zweig, in: Hamann, 1969, p. 43).

Apart from the complex image of woman-truth, it is worth noting the distinction between philosopher and poet, lover and visionary, the demand that false gods be mocked, and the allusion to the myth about the Tower of Babel.

4.1. Adaptation from *The Robbed Fable*
(Die beraubte Fabel) by Lichtwer

This image of truth is an adaptation from *The Robbed Fable* (*Die beraubte Fabel*) by Magnus Gottfried Lichtwer from 1748 (Lichtwer, 1978, p. 52). Lichtwer tells the story of the woman-fable, the goddess of all poets, who sets off on a journey during which she is assaulted by robbers. Since her purse is empty, they demand that she give them all her clothes. The fable therefore takes off her dresses, one by one, seemingly endlessly. When she stands before them naked, they realize she is the truth and decide to give back her clothes, as they cannot bear to see the naked truth.

Lichtwer's poem is an allegorical depiction of the literary genre of the fable, which itself is allegorical in character. Lichtwer seems to claim that in the fable, different kinds of clothing and animal disguises illustrate some general universal moral truth that is not necessarily expressed explicitly in the moral within

the fable's text. This moral truth can be expressed otherwise, not only by means of allegory but also directly, by means of general terms as well. The latter way, however, misses the didactic aim of the fable – the naked truth is unbearable, no one can bear to look at it, no one is able to absorb it.

If we take *The Robbed Fable* as the source of a metaphor outlining a certain concept of truth, then in Lichtwer's fable we will see a humorous expression of the idea that Blumenberg[4] calls modern skepticism as to the purpose of attaining the truth. It is not the attainability of truth that is questioned here; reaching the naked truth is seen as possible. Lichtwer's metaphor doubts, however, if it is sensible to uncover the truth in the face of the fact that such truth is unbearable. The only assimilable, digestible truth is truth in disguise; only such a form enables us to live (Blumenberg, 2010, p. 48).

The changes Hamann introduced to Lichtwer's fable are not random, and according to metaphorology one can see an essential shift in thinking about truth in them. Hamann's adaptation introduces the following innovations: truth's identity is known from the very beginning of the story, not after the disrobing. Hamann does not speak about the fable which is truth in disguise, but about truth itself, who is a woman in many layers of dresses. After the disrobing, the robbers see not a naked body whose existence they doubt, but a ghost, an apparition, a fearful bodiless phantom. The layers of clothing do not cover the identity of truth, they are not a disguise, but have a protective function. They protect the truth from violence, and they also hide the fearful "truth about truth," namely that truth does not have a body. The dresses are not something secondary and external to the truth, but are constitutive for it – when they are gone, we face a different order of reality: a ghost or a phantom. It is not clear whether Hamann's highway robbers are interested in robbery; rather, he suggests they have designs on their victim's virtue and try to win her by force, which turns out to be impossible because she has no body.

---

[4] In the chapter "Metaphorics of the 'Naked' Truth" (Blumenberg, 2010, pp. 40–51) in the book *Paradigms for a Metaphorology*, Blumenberg does not mention Hamann or Lichtwer, but they can both be interpreted within the framework of Blumenberg's analysis, and they can also enrich this framework with some additional nuances of the historical change of metaphors of 'naked' and 'covered' truth.

4.2. The one and only truth

Hamann believes in the existence of truth. He claims there is only one truth, that connected to divine reality, which manifests itself in events of the history of salvation (the creation, incarnation and redemption)[5]. In *Biblical Reflections of a Christian* (*Biblische Betrachtungen eines Christen)* Hamann writes: "There is one and only truth, though it has countless comparisons and expressions" (N I, 47, 19). Similarly, he claims in his correspondence with Lindner from May 5, 1761: "*Opinions* are only a *vehiculum* of truth, not truth itself" (ZH, II, 84, 24f). These two quotations show how Hamann understands the expressibility of truth. It is expressible, yet it can never be fully expressed, because its expressions are limited to their spatiotemporal dimensions. The vehicles of truth make the experience of truth possible, but they can never fully capture it.

Truth cannot be defined in philosophical terminology once and for all, it cannot be separated from other concepts. To Hamann, such a desire is a misconception of truth, and the very question of Pontius Pilate, "What is truth?" is wrongly expressed and testifies to an improper attitude toward the truth (Stünkel, 2007, pp. 159–160). Truth is not an abstract concept but something definitely different. One listens to the truth, testifies to the truth, strives for truth or lives according to it. Truth is what God reveals about Himself to humans.

Hamann's model of how truth manifests itself has a distinct theological dimension: It is based on the mechanism of condescension and self-emptying of God (see: Reuter, 2005; Terezakis, 2012). God Himself is the truth and allows humans to get to know him. Condescension is a communicational act through which God speaks to humans. But for the message to be available to the recipient, God needs to adapt Himself to human abilities by assuming a sensorial form that can be perceived by humans.

Divine truth reveals itself and allows itself to be received by humans through the incarnation of Christ, the creation, Holy

---

[5] In *Socratic Memorabilia* he writes: "If it is true, that God Himself, as the **good testimony** says that he made before Pilate; if it is true, I say, that God Himself **for this** became human and **for this** came into the world, **that he might witness to the truth**: then it requires no omniscience to see in advance that he would not depart from the world as well as Socrates, but rather would die a more shameful and cruel death" (N, II, 82, 14–21, transl. G. Griffith-Dickson, 1995, p. 400).

Scripture, history, and human words. In fact every sensorially perceivable manifestation of reality can be treated as a message from God. In each case, the translation of invisible reality into that which is concrete, sensorial and historical, is different. God's truth is even embodied in writings of adversaries of religion, whom Hamann sees as prophets "without awareness or desire" (ZH, I, 380, 7–8).

4.3. Incarnation as a matrix for the expression of truth

Incarnation is a special kind of God's condescension. Thanks to incarnation, invisible divinity assumes visible, human shape. The second person of the Trinity limits himself and assumes a human body and nature. This image of truth manifesting itself in the act of God's self-limitation describes the situation of the transmission of God's truth in the Bible as well as the expression of the author's own thoughts.

> The invisible nature *[Wesen]* of our mind *[Seele]* is revealed by words ... Is it not humiliating for thought that, as it were, [spirit and body] are incapable of visibility apart from the crude clothing of arbitrary signs; and what a proof of divine omnipotence – and humility – that He could and would inspire the conceptions of men with the depths of His secrets, the treasure of His wisdom, in such gibberish, such confused tongues (ZH, I, 393, 28–394,3; transl. A. Anderson, 1969, p. 253).

Just like in the metaphor from Hamann's letter to Kant, in the very beginning truth is something bodiless, spiritual and invisible, irrespective of whether it is God's *secret and wisdom* or human *mind and thought*. It receives visible shape through linguistic expression, which however is *crude clothing* – always insufficient to express the whole truth. Therefore the act of speaking the truth must be constantly repeated in changing existential and historical circumstances (see: Russ, 2013, p. 19).

Elsewhere Hamann writes:

> The truth is like a seed, to which man gives body as he wants, and this body of truth again receives clothing through an expression, according to everyone's taste or according to the principles of fashion (ZH, I, 335, 28f).

This quotation comes from a letter to Lindner dated June 1, 1759 containing an important reflection on how truth is expressed in both God's salvation acts and a writer's creative work. Hamann mentions a fragment of Augustine's *Confessions* which refers to ambiguous biblical language that constitutes a specially designed form of transmission of God's truth, which can be interpreted in manifold ways, thanks to which different people in different times can receive it in a suitable way (Augustine, *Confessions,* XII, 26/31; see: Beech, 2006, pp. 33–36; Anderson, 1969).

Hamann refers to Augustine in polemic with Chladenius in the context of considering the style of the apologetic writings of a Christian author. Chladenius prefers the ideal of clarity and simplicity in interpreting the Bible as well as in apologetic writing. Hamann, on the contrary, refers to Augustine's ideal of ambiguous terseness of biblical text as the model for his own writing.

> The concepts that Augustine accepts to some extent contradict the first rules which we used to consider good style. Instead he assumes that truth could exist in the great multitude of opinions on one and the same thing, therefore he would like to write in such a way that those who have an understanding of God's creative power, in quamlibet sententiam cogitando venisset, eam non praetermissam in paucis verbis tui famuli reperirent et si alius aliam vidisset in luce veritatis nec ipsa in iisdem verbis intelligenda deesset … [those who can already grasp it, whatever true meaning they had reached by reasoning they would find it not neglected in your servant's words. And if someone else had seen another meaning in the light of truth, that too would not fail to be comprehensible in the same term] (ZH, I, 335, 9–17; transl. C.J.-B. Hammond, 2016, p. 319).

The fact that Holy Scripture can be read, understood and related to personal life in all ages does not stem from the eternal, timeless truths it conveys, but from the multilayered language which can be interpreted in many ways, depending on the context, without losing its veracity.

Truth is not available to humans as the one and only. Timeless and abstract truth is a ghost, a phantom. It is available only in translation into its sensorial, concrete realizations – dresses that depend on the time and place. When Hamann speaks about the availability of truth through its *vehicula:* "dresses," "comparisons," "expressions," "opinions," he does not only mean the "digestibility" or assimilability of truth as in Lichtwer (see: Blumenberg,

2010, p. 48), but first and foremost indicates truth's fundamental and constitutive features.

One of the most important characteristics of truth is its temporality and historicity. Hamann expresses this with an ancient phrase found in Bacon: Truth is "the daughter of time" (ZH, VI, 162,26; cited in: Bayer, 1991, p. 97).

Truth is available to humans in its historical manifestations, i.e. incarnated, clothed in a garment that is concrete, figurative and individual. Hamann presents the condition of humans condemned to prospective access to truth through the biblical myth about the Tower of Babel. The confusion, fragmentation and multiplication of languages is not only punishment for attempted self-deification, it is motivated by God's love, friendship and concern for humanity. Philosophers-builders of systems search for truth improperly, while the experience of truth is possible through faith and grace, not by means of self-reliant reasoning. Only God is able to reverse the consequences of speech confusion. And He does it not by reducing all languages into one, but by enabling people to convey the Good News in many languages and in many different ways, as it happened on the day of the Pentecost.

5. Authorship as serving the truth

The image of the truth in many garments enables us to distinguish two types of relations toward the truth. The first involves using violence for the acquisition of truth. This can be understood as the conceptual violence of philosophers expressed in the struggle to purify truth of its contingent characteristics, stripping her of clothes by means of abstraction. For the robbers from the metaphor-fable, the process of stripping truth of her dresses has to culminate in the terrifying discovery that behind the concept of truth purified from all contingency there is literally nothing.

Philosophical violence is also presented through a divorce metaphor: "The philosophers have always given truth a bill of divorce, by separating what nature has joined together" (N, III, 40, 3–5; transl. G. Griffith-Dickson, 1995, p. 480). Ultimately, though, this aim appears to be unattainable, and Hamann sees its motives as a desire to dominate and control, as well as the sin of pride and idolatry, namely worshipping a self-constructed projection of truth instead of the hidden truth of God.

The other attitude toward truth can be called the attitude of "poets, lovers, and visionaries." It is to be implemented through love, sacrifice, serving and testifying to the truth. Hamann sees truth as his beloved, alluding to the bride from *The Song of Songs*: "Truth is my girl, black but charming, like the tents of Kedar, like the carpets of Solomon" (N II, 103). Hamann writes about serving the truth in his *Socratic Memorabilia*, where he makes Socrates, "the prophet of an unknown God," an example of the proper relation to the truth. Such an attitude should be based on respect and love for the truth, which does not cause a desire to conquer the naked truth but, on the contrary, means cultivating its manifold incarnations through poetry and prophecy.

This is how Hamann understands the vocation of a writer when he includes an anecdote about Socrates the sculptor in *Socratic Memorabilia*, where the notion of clothing of thoughts reappears:

> Socrates presumably was a sculptor, because his father was one. One may deduce that he did not remain mediocre in this art, for his three statues of the Graces were erected in Athens. At one time, it was customary to clothe these goddesses; Socrates followed this old-fashioned practice and his Graces contradict the custom of the contemporary system of the gods and the fine arts based on them. How Socrates came to this revival, whether it was a contribution of his Genius, or a vain desire to distinguish his work from others, or the simplicity of a natural modesty which must have seemed extraordinary to a pious Athenian; I do not know (N, II, 66, 1–11; transl. G. Griffith-Dickson, p. 385).

Referring to this passage in *Clouds* (*Wolken*), Hamann wrote: "The reflections on the statues of the Graces contain a defense of this clothing, which is proper only for chimerical ideas" (N II 92, 21–23). From the beginning of his authorship Hamann defended metaphorical, pictorial language as a fundament of human experience. It is not a rhetorical addition to utterances, a form unrelated to the content or a poetical decoration. It is something more: a way of thinking that does justice to the truth.

6. Conclusion

The metaphor of truth in many garments is not only emblematic of Hamann's concept of truth but is a key point binding

other subjects he considered crucial, such as language, the condescension of God, and the vocation of an author.

It is worth noting that Hamann touches on the same topics that are within Blumenberg's primary interest: The metaphor of truth is not only a pre-conceptual sign of change in thinking about truth which occurred in the Enlightenment and to which Blumenberg refers, without mentioning Hamann, as follows:

> The discovery of history within the Enlightenment and against its semantic current is the discovery of the illusion of the 'naked truth' or the illusoriness of nakedness. It is at once the negation of the metaphor and its renewal in a different direction: the 'disguises' of truth now no longer issue from the poetic imagination and the need for rhetorical ornamentation—they do not represent fashion accessories, so to speak, that could be cast off without further ado—but are themselves constitutive of the way in which truth manifests itself (Blumenberg, 2010, p. 48).

From another point of view, the analyzed metaphor of truth at the same time forms a defense and indicates the necessary presence of metaphorical language in philosophy, its irreducibility and non-substitutability. It is worth noting that Hamann expresses the "truth about truth" by means of many metaphors and images; it is possible to build more than one interpretation for each of them. Such diversity and ambiguity is an attempt at a dynamic study of truth, at grasping it from different standpoints at the same time. This examination is aimed at evoking all these images in readers, in order to let them experience love toward the truth and draw them into serving it.

References

Alexander, W.M. (1966). *Johann Georg Hamann. Philosophy and Faith*. The Hague: Martinus Nijhoff.

Anderson, A. (1969). Philosophical obscurantism: Prolegomena to Hamann's views on language. *Harvard Theological Review, 62*, 247–274.

Augustine (2016). *Confessions. Book 9–13*, transl. C.J.-B. Hammond. Cambridge, MA: Harvard University Press.

Bayer, O. (1991). *Autorität und Kritik. Zur Hermeneutik und Wissenschaftstheorie*. Tübingen: Mohr Siebeck.

Beech, T. (2002). *Hamann's Prophetic Mission. A Genetic Study of Three Late Works against the Enlightenment*. London: Maney Publishing for the Modern Humanities Research Association.

Blumenberg, H. (2010). *Paradigms for a Metaphorology*, transl. R. Savage. New York: Cornell University Press and Cornell University Library.

Griffith-Dickson, G. (1995). *Johann Georg Hamann's Relational Metacriticism*. Berlin: Walter de Gruyter & Co.

Hamann, J.G. (1949–1957). *Sämtliche Werke*. 1–6. Josef Nadler (Ed.). Vienna: Herder (quoted as N).

Hamann, J.G. (1955–1979). *Johann Georg Hamann Briefwechsel*, vol. 1–7, W. Ziesemer, A. Henkel (Eds.). Frankfurt a. Main: Insel-Verlag (quoted as ZH).

Hamann, J.G. (1967). Letter from J.G. Hamann, July 27, 1759. In: A. Zweig (Ed.), *Kant. Philosophical Correspondence 1759–1799*. Chicago: University of Chicago Press.

Hamann, J.G. (1996). *Letter to Christian Jacob Kraus (18 December 1784)*, transl. G. Green. In: J. Schmidt (Ed.), *What Is Enlightenment? Eighteenth-Century Answers and Twentieth-Century Questions* (pp. 145–154). Berkeley: University of California Press.

Lichtwer, M.G. (1978). Die beraubte Fabel. In: E. Leibfried, J.M. Werle (Eds.), *Texte zur Theorie der Fabel* (p. 52). Berlin: Springer.

Moustakas, U. (2003). *Urkunde und Experiment: Neuzeitliche Naturwissenschaft im Horizont einer hermeneutischen Theologie der Schöpfung bei Johann Georg Hamann*. Berlin: Walter de Gruyter.

Regier, A. (2016). Johann Georg Hamann. Metacritique and Poesis in Counter-Enlightenment. In: P. Hamilton (Ed.), *The Oxford Handbook of European Romanticism. First edition* (pp. 165–183). Oxford: Oxford University Press.

Reuter, Ch. (2005). *Autorschaft als Kondeszendenz. Johann Georg Hamanns erlesene Dialogizität*. Berlin: Walter de Gruyter.

Russ, A.R. (2013). *The Illusion of History: Time and the Radical Political Imagination*. Washington: CUA Press.

Stünkel, K.M. (2007). Ästhetische Geologie. Die Frage nach der Wahrheit bei Johann Georg Hamann. *Neue Zeitschrift für Systematische Theologie Und Religionsphilosophie, 49* (2), 156–182.

Terezakis, K. (2012). Is theology possible after Hamann? In: L.M. Anderson (Ed.), *Hamann and the Tradition* (pp. 182–198). Evanston: Northwestern University Press.

Marta Baron-Milian
ORCID 0000-0002-5430-4339
University of Silesia in Katowice

Worthless yet Priceless: The Truths and Economics of Poetry

Abstract

The essay tries to capture the epistemological status of poetry from the perspective of questions posed by the economy of literature. Selected theories of poetry – those of Jochen Hörisch, Viktor B. Shklovsky, Jean Baudrillard and Franco Berardi – can be regarded as a kind of "economics of poetry" due to their proposed treatment of the properties of the poetic medium from an economic perspective. In each of these theories, poetry is defined as uneconomical and useless, which, paradoxically, becomes its fundamental condition for existence as a place for breaking the linguistic circulation of signs, a point of resistance against conventionalized communication and automatization of perception, and thus a medium offering unique knowledge about reality.

Keywords: poetry, truth, economy of literature

> "Confusion to the memory of Newton …
> because he destroyed the poetry of the rainbow"
> John Keats

> "Cursed be the glare of apparition
> That on the finer sense intrudes!"
> Johann Wolfgang Goethe

Sometimes little can be seen – even though it might be illuminated, hidden in plain sight. A flash of light can be blinding, causing our eyes to narrow in the bright sunlight. Introducing readers into the sphere of disturbing oxymora and confronting a peculiar cognitive crisis, in the poem (or, rather, prose poetry) *Darkness*, Stanisław Barańczak wrote: "One needs the power of darkness to know more clearly" (Barańczak, 2007, p. 40). While both logic and knowledge point to a different outcome,

for Barańczak brightness is not enough to know clearly – here the condition of cognition is "the power of darkness." It is not about "seeing more clearly," however, even though the syntax and phraseology suggest it to be so, but more about "knowing more clearly," even if this seems to be a spelling mistake. If we take the path of truth to be the path of reference, confirming congruence between words and reality, the quote from *Darkness* erases all addresses that direct words to things, pointing rather to the devious cognitive paths of aporias, errors, and paradoxes. A fiasco from the point of view of logic and reference turns out to be a victory in a different space – that of figuration and contrasting perception. A supposed cognitive defeat, along with the affirmation of an erroneous trope, is, in some sense, a victory. The piled-up paradoxes found in Barańczak temporize with the truth, rhetorically repealing the necessity of verification and, at the same time, postponing accusations of falsehood. It is as if poetry had its own answer to a question from which even logic backs away: "One needs the power of darkness to know more clearly."

Further complicating the meanings of brightness and darkness in poetic cognition, let us begin by dispersing the light. It would be difficult to conjure up a better tale illuminating all the tensions outlined here than that of the scientific and poetic case of the rainbow[1]. Through experiments with the prism, Isaac Newton proved that white light splits into the full-color spectrum. Publishing his famous *Opticks* in 1704, Newton could not have imagined the gravity of his discovery for poets. To enlightened society, Newton's "untangling" of the threads of light became a fascinating scientific axiom that, in a sense, was poetic in itself, and, aided by metaphors, could additionally become sensual, dramatic and charming. "Did ever poet image ought so fair, dreaming in whispering groves, by the hoarse brook … How just, how beauteous the refractive law," wrote James Thomson in 1727 in the poem *To the Memory of Sir Isaac Newton* (Abrams, 2012, p. 334). Yet this poetic fascination with science was only momentary, it waned at the turn of the 18th and 19th centuries, setting a permanent opposition between poetic truth that

---

[1] The following story of the Romantics' rainbows was inspired by the monograph *The Mirror and the Lamp: Romantic Theory and the Critical Tradition* by M.H. Abrams (Abrams, 2003).

(fortunately) was impossible to verify and verifiable scientific judgment. And so began the poetic mourning of the rainbow, with John Keats' funeral verse: "There was an awful rainbow once in heaven:/ We know her woof, her texture; she is given/ In the dull catalogue of common things./ Philosophy will clip an Angel's wings,/ Conquer all mysteries by rule and line,/ Empty the haunted air, and gnomed mine —/ Unweave a rainbow, as it erewhile made/ The tender-person'd Lamia melt into a shade" (Abrams, p. 337). Poetic condemnation of scientific scrutiny will have no end...

Thus the arc of the rainbow binds together the problem of truth and poetry, but there still remains the final element of this essay: economics. Could it be considered a lucky coincidence that a leprechaun, made famous by Irish legends, hides his treasure – a big pot of gold – at the end of the rainbow?

Brighter, darker...

> "All poetry is misrepresentation"
> Jeremy Bentham

To consider only the 20th century, literary work was not only a separate body from scientific work, but was often also in opposition to it. Danuta Ulicka enumerates such approaches by a number of scholars: "in 1926 by Richards, 1931 – Ingarden, 1942 – Wellek and Warren, in 1949 – by Kayser, 1957 – Frye, ... in 1972 by Genette and 1975 by Culler" (Ulicka, 2006, p. 10). Nevertheless, literary works continued to emit (peculiar) "cognitive toxins" (ibid.). These toxins turned out to be so mysterious and stimulated the artistic, literary and philosophical imagination to such an extent that they were scrutinized over and over again. Because of this, nowadays it is nearly impossible to briefly or selectively illustrate the history of reflection on the epistemological status, truths, wisdom or cognitive functions encapsulated in poetry, for these concepts are inextricably linked with a number of others that are related: mimesis, representation, reference, realism, etc...

In his most famous lecture, about placing poetics within the linguistic model, which is connected with the problem of the arbitrary nature of the sign as described by Ferdinand de Saussure, Roman Jakobson subordinated the poetic function to the

referential function, but also warned that "Any attempt to reduce the sphere of the poetic function to poetry or to confine poetry to the poetic function would be a delusive oversimplification" (Jakobson, 2006, p. 247). And so, researchers began using literary terms as if literature were akin to scientific knowledge, never really attempting to determine the original character of literary cognition. Since defining it was an impossible task, metaphors were mainly used when trying to express the cognitive function of literature. That is why the archive of metaphors used in an attempt to grasp the relationship between a literary text and reality, literary fiction and truth, or between the word and the thing being represented by it, is extremely vast. We are condemned to wandering the treacherous corridors of this archive because once we choose a certain key to unlock one of the infinite doors, we quickly realize that we have been led astray and are walking the archive's corridors in circles again. The extent of the problem (and the trouble with the keys) is shown even by just a few of the paradoxes found in reflection on the relationship between literature and the world. As Antoine Compagnon aptly but ironically points out: "In Plato's *Republic*, mimesis is subversive, it threatens the social bond and poets must be banished from the City ... . At the other extreme, for Barthes, mimesis is repressive, it consolidates the social bond because it is allied with ideology (the *doxa*), which it serves as an instrument. Is mimesis subversive or repressive? Since it can take on such disparate qualifiers, it is probably not the same notion; from Plato to Barthes, it has been thoroughly reversed, but between the two, from Aristotle to Auerbach, it was not seen as anything harmful" (Compagnon, 2004, p. 70).

It is not surprising that this whole collection of metaphors from literature and literary studies, attempting to capture the epistemological status of a literary work, was described by Rita Felski as "a chronicle of outgrown errors" (Felski, 2016, p. 92). In this collection of metaphors, we shall find, above all else, probabilities and fictions, appearances and shadows, illusions and delusions, fakes and counterfeits, imitations and copies, deformations and distortions, substitutes and equivalents. From the point of view of the critique of literature as not respecting the truth, which goes back as far as Plato, a literary work is "a sham, a shameless un-truth, this failure of knowledge drawing all kinds of calamitous consequences in its wake" (ibid., p. 87). And yet,

another sphere of metaphors disperses the darkness of the metaphor archive: metaphors in which literature is perceived as light or a lamp[2] revealing, thanks to its unique quality, what has yet not been seen, illuminating and exposing the real or different nature of things. Nevertheless, the power of literature is at times compared to that befitting "second nature," i.e. presenting the literary work as a microcosm governed by its own laws or, more often, a "heterocosm" in which the poet's potential is equaled to God the creator (Abrams, 2012, p. 298). The darkness of the metaphor archive will also be illuminated by the cognitive optimism of several other well-known glass objects, such as a mirror reflecting reality, or a window through which one can look at the world outside. Both metaphors have been criticized, however, for the same reason as the whole theory of reflection (as developed, for example, by György Lukács): for the illusion of transparency; transparency that ignores the specifics of linguistic mediation, or, as Michel Foucault puts it in *Words and Things*, for the utopian belief that things can be named without any interference (Foucault, 2005, p. 133). In the meantime, what we are used to referring to as reality, in literature appears only as the "reality effect" (Barthes, 1999, p. 118) or the "code of representation" (ibid.), which in turn only exists as a "perspective of quotations" (ibid. p. 55). At this point, reference is being replaced by intertextuality. However, once we dig further into the archive, we arrive at the map section, illustrating researchers' love for visual analogies, showing cartography to be still in fashion, as in the cognitive mapping of Frederic Jameson (Jameson, 2009), the maps and graphs of Franco Moretti (Moretti, 2016), to cite just a few examples. Further, we find a room where texts can be recognized as symptoms, and some researchers (e.g. following Althusser) agree that literature, while unable to convey true knowledge, "can attain a critical distance from the everyday work of ideology by rendering it in aesthetic form, thereby exposing its repressed or excluded meanings" (Felski, p. 90). The relationship between reality and art is no longer derived from the notion of similarity but is seen through the prism of suppressed causality,

---

[2] Among many other sources, the most complete and interesting analysis of these metaphors is to be found in the aforementioned monograph by Abrams (Abrams, 2003). An example of a synthetic approach to the problem in Polish literary studies can be found in Anna Krajewska's essay, "Light as an Epistemological Metaphor" (Krajewska, 2006).

referring to social conditions. But following the tropes of subsequent metaphors, one cannot overlook the tropes themselves, for literature is at times perceived as a trail or trope of reality (Nycz, 2001), or a space of unexpected epiphanies (ibid.). Finally, the archive of metaphors (guiding us toward the truth of literature) draws attention to the very concept of metaphor as an essential tool of literary cognition[3]. And so, a metaphor does not falsify nor verify, but changes our point of view and way of looking at things, allowing us to scrutinize them more intently, obscuring the general picture so we might "know more clearly." This rather long list serves not as an attempt to exhaust the resources of key tropes that appear in reflection on literary cognition. Rather, as is the case with all enumerations, it attempts to showcase a small but most representative part instead of the overwhelming whole[4].

Criticism of poetry through its cognitive function reached its peak in positivism, where recognition of the fact that a poetic work does not generate knowledge became the basis for speaking about it in terms of uselessness. Jeremy Bentham judges poetry extremely harshly, posing two simple (and "lethal") questions: "Is it useful?" and "Is it true to reality?" (Abrams, 2012, p. 329). John Locke's *Some Thoughts Concerning Education* leaves no shadow of a doubt as to the answer, stating that both poetry and gambling "which usually go together, are alike in this too, that they seldom bring any advantage but to those who have nothing else to live on" (Abrams, 2012, p. 329). To positivists, poetry was only useful so far as it offered the possibility of pleasure and entertainment. In a cognitive sense, it was often considered not only useless but also harmful – falsifying reality.

---

[3] Not only in the literary sense, of course, but also the philosophical. The most famous lectures can be found in the writings of Friedrich Nietzsche (Nietzsche, 1993), Jacques Derrida (Derrida, 2002) and Hans Blumenberg (Blumenberg, 2017).

[4] An important synthetic and analytical approach to the problem of the "truth of poetry" can be found in the book *Truth in Literature* (edited by Andrzej Tyszczyk, Jarosław Borowski and Ireneusz Piekarski). Especially vital to the problem would be, among other examples, Janusz Misiewicz's essay "The Truth of Poetry," Bernadetta Kuczera-Chachulska's "About Lyricism and the 'Truth of Life'," and Witold Sadowski's "Tone and Truth in Works Written in Verse." A series of important reflections on the epistemological status of poetry is also brought by essays collected in the book *Literature and Knowledge* (edited by Włodzimierz Bolecki and Elżbieta Dąbrowska), to which I refer in this essay a number of times.

Although it can be said that these qualities mark literature and art in general, it is poetry that was denied economic functionality to the greatest extent due to, among other things, its extremely confused relationship with truth and reality. From a scientific point of view, poetry (much more than other literary genres) appears to be useless, and from an economic standpoint – superfluous and unnecessary and, therefore, simply uneconomical. Since poetic words cannot be simply exchanged for wares, and all referential functions are highly problematic, the losses definitely outweigh the cognitive profits that one could gain from a search for "poetic truth." Cognitive worthlessness and the logic of unprofitability surrounding poetry situates it at the outskirts of what we could call the economics of communication, representation, and knowledge. While the above seems like a debate over an issue already resolved (at least since the 19th century), it does draw attention to the peculiar relationship between thinking about poetry in economic terms[5] and discovering or differently understanding its cognitive potential. This kind of poetic uselessness and lack of economy can be treated not as the offensive diagnosis of superficiality, but as a peculiar kind of pricelessness, establishing poetry's unique position in the market of signs.

It is exactly this combination of worthlessness and pricelessness that forms the basis of a number of theories of poetry, developed in different contexts (philosophical, artistic, historical, political, etc.). Tracing various theories of poetry, as it turns out, we still encounter new ones developed and described in the rather unfavorable language of economics. Oftentimes theories of poetry, poetic program manifestos, and meta-artistic statements formulated by both literary scholars, philosophers, and artists, whether of a literary or historical or critical nature, reveal a unique tendency to conceptualize poetry and characterize the poetic medium in economic terms, using economic concepts, metaphors and theories. These theories take the worthlessness or uneconomical character of the poetic form and turn it around to show that through this apparent loss, poetry offers the possibility of unique profits of a cognitive and critical nature.

---

[5] One of the most interesting examples of such thinking is given by Richard Sieburth in his analysis of Ezra Pound's work, calling Pound's oeuvre "economics of poetry," "poetry of economics" or "poetics of money" (Sieburth, 1995).

It seems, however, that this important issue (combining numerous theories of poetry) does not find much attention within the economy of literature[6] or in what we could call economic literary criticism[7]. The economy of literature began to develop dynamically in the 1980s, dispersing in very different directions, yet in its majority focusing on the study of narrative genres, with the novel playing a key role. Therefore, we are dealing with the marginalization of poetry in studies on literary cognitive function, specific literary knowledge, literature as social practice or the area at the intersection of various external political, economic and social determining factors – all being fields from the economy of literature. This omission seems significant, denying poetic texts not only cognitive value and the role of purveyors of knowledge about reality, but also transferring them to the sphere of utopian autonomy, detached from such texts' heteronomous nature. Thus, not only poetry's cognitive and critical possibilities are questioned, but also its potential for political influence and, consequently, its social significance. Meanwhile, asked from the point of view of economics, questions about the transaction between poems and the world seem to open not only the possibility of a different approach to the problem of poetic cognition but a different placement of the theory of poetry as well.

Disrupted circulation of words

"Who wants to be Linnaeus at the heart of the sun"
Andrzej Sosnowski

Perhaps it is no coincidence that one of the most coherent projects combining the problems of truth and economics of poetry, authored by German historian of literature and philosopher Jochen Hörisch[8], was based primarily on an analysis of

---

[6] This concept appears in the canon for this research area, Marc Shell's *The Economy of Literature* (Shell, 1979).

[7] American research on the economics of literature was conducted under the project name of New Economic Criticism in the 1990s. The name was also used as the title of their monograph: *The New Economic Criticism: Studies at the Interface of Literature and Economics* (Osteen, Woodmansee, Eds., 1999).

[8] Hörisch has not been very popular in Polish literary studies thus far. Paweł Tomczok, in his monograph *Literary Capitalism* (Tomczok, 2018, pp. 67–76), comments more broadly on the theory of media developed by Hörisch.

the work of the great researcher, admirer and "theoretician" of the rainbow – Johann Wolfgang Goethe. With dedication comparable only to his work on *Faust*, Goethe devoted years to his *Theory of Colors*, published in 1810. Experimenting with a prism, Goethe distanced himself from Newton's theory of the splitting of white light. In his peculiar and romantic "phenomenology" or "psychology" of colors, he tried to show that they have their source in the interaction of light and darkness. Although Hörisch's interest in *Faust* is more concerned with metaphors related to money rather than rainbows, in the Polish translation of Goethe's drama we find a fragment in which both these poetic delusions, monetary and rainbow-colored, are combined, and are maybe equally important: "Cursed be the glare of [rainbow] apparition/ That on the finer sense intrudes! ... Cursed Mammon be, when he with treasures/ To restless action spurs our fate!" (Goethe, 2018, p. 38). This commonly recognized quality of poetry and economics became the starting point for Hörisch, author of the famous trilogy dealing with media history and theory (of which *Heads or Tails: The Poetics of Money* is a part).

Hörisch's approach should be presented here first, for it directly combines economics of poetry with the problem of literary *episteme*. He places this relationship at the center of his reflection, giving the poetic form historical character and an important social function, directly resulting from the cognitive and ontosemiological values of the poetic medium. Furthermore, the basis for the concept presented in *The Poetics of Money* is recognizing that literature, especially poetry, is functionally unnecessary and useless. Hörisch turns the subject around, showing that this superfluity neither eliminates the cognitive possibilities of literature nor results from them, but is an essential condition for the production of specifically literary knowledge. As we read in *The Poetics of Money*:

> Belles lettres is always surrounded by the suspicion of being functionally superfluous – 'I am that spendthrift, poetry.' That which is superfluous lives on excess and can therefore afford what others must renounce. Literature began to perceive itself early, and has continued to do so with increasing intensity since the sixteenth century, as the medium that does *not* require cover and is therefore all the more qualified to observe these problems. Literature, after all, does not even claim to validate its statements. Literature is fiction, which means, precisely, that it does not have to be covered by actual events or realities (Hörisch, 2000, p. 17).

The lack of coverage and confirmation of literary statements in reality gives literature the opportunity to present alternative ways of perception. It thus provides "alternative versions and interpretations" of reality (ibid.). As Hörisch notes, literature as an ontosemiological medium enables social synthesis, giving rise to "a sociocultural tapestry that provides every individual event with an orienting framework" (ibid., p. 23). Literature, being a luxury, can afford unwanted commentary that other discourses cannot: "… literature interferes … literature irritates accepted codes, literature interrupts and disturbs communications and common figures of speech" (ibid., p. 27). Such a disruption of communications, found especially in poetry, is the main, under-appreciated competence of literature, inextricably linked to its cognitive function. This quality is fully evident in the perception of the history of literature as a history of problems, in which the literary medium is a peculiar resource of knowledge unavailable to the discourses of other knowledge systems, as it focuses on differences in presenting certain factual issues. Hörisch defines this peculiar form of literary knowledge after Walter Benjamin, to whom its indicator would be the intertwining of a given work's philosophical content and its truth.

Hörisch's literary history project thus enhances the cognitive potential of poetry by displaying the heteronomous nature of art, at the same time situating itself on the outskirts of thinking in terms of textual autonomy. In order to capture the nature of poetic transactions with the world (being at the same time an argument for poetry's combination of worthlessness/pricelessness analyzed in this essay), it is undoubtedly necessary to shift to the level of signs.

Thinking in terms of exchange and the "market of signs" is a kind of "ground zero" in understanding literature and poetry in economic terms, in close connection with their cognitive function. The key belief here is that all semiotic theories based on the notions of equivalence, value, or sign exchange, have their common genealogy in the economic discourse. Reflecting on the relations between economics and representation, Michał Paweł Markowski notes: "Just as there is no pre-monetary economics, there is no pre-sign semantics, which means that meaning, just like monetary value, is created only because that which is individual can be exchanged for what is general. This process of exchange can be defined in various ways: as generalization,

abstraction, or symbolization, and it would be right to introduce the category of representation here as an integrative category" (Markowski, 2004, pp. 12–13). However, not only economics and semiotics are combined in the simplest definition of representation, "something exchanged for something else," but also thinking itself, based on the transition "from sensual contact with what is individual to abstract knowledge, using mediatizing categories" (Markowski, 2004, p. 13). Rhetoric, epistemology, and economics come together on the horizon of representation, understood as equivalence. Markowski also justifies this shared genealogy etymologically: *repraesentatio* can be a concept in the mind as well as "hard cash" that can be paid immediately (Markowski, 2004, p. 13).

Just as money is exchanged for commodities because of its shared value, so are words exchanged for things they stand for through common meaning. If we take the economics of communication as the starting point, then the perfect implementation of this kind of (rather reductionist) economics of literature would be realism. Realism maintains its own illusion of reference: the illusion of adequately exchanging things for words, being a medium of the transparent, pre-linguistic character of reality whose proper representation is used by people aware of themselves and the world around them. Markowski uses this example as a starting point for his project on modern literature as being uneconomical, defying economics (Markowski, 2007, p. 60), with modern poetry as an absolute model of "non-exchangeable speech" (ibid., p. 97). Modern literature manifests the opacity of its own medium and recognizes itself as a place of poetic resistance to everyday language ruled by the principles of the economics of communication. According to Markowski, this is the point where the critical potential of modern literature is revealed – dealing with the crisis of representation, undermining the traditional definitions of mimesis, realism, references, truth and fiction, medium and representation. Poetry is paradigmatic for this economics-related relationship between realism and modern literature, with critics thinking of it as non-exchangeable[9].

---

[9] One of the examples for Markowski is the critical discussion between Ignacy Fik and Stefan Napierski about Bruno Schulz's prose, in which they identify ostentatiously poetic and uneconomical qualities, making claims to literary truth and representation, and accuse his prose of abnormality, perversity, horrors and, especially, uselessness.

Although similar thinking in terms of communicative exchange and poetic non-exchangeability could be traced back to antiquity, it is in 20th-century theories of poetry that this approach makes a significant impact, Viktor B. Shklovsky's formalist dissertation *Art as Technique* being a primary example defining the functions of poetry using the language of economics. Shklovsky begins with a challenging concept of economy: "Poetry is a special way of thinking; it is, precisely, a way of thinking in images, a way which permits what is generally called 'economy of mental effort' ... Aesthetic feeling is the result of this economy" (Szkłowski, 1986, p. 13). Following various theoretical tropes, Shklovsky presents the "general law of saving spiritual strength" or "the principle of the economy of creative effort" (ibid.) – principles on which the economics of poetry is based, in the opinion of researchers mentioned in this essay. Shklovsky's goal is to show the misunderstanding lying at the heart of this perception of the economics of poetry, i.e. based on the principle of economy, which is a solid principle for communication/practical language, but when extended to poetic language serves as evidence of complete ignorance of its laws. Instead of economy then, we should talk about "the laws of expenditure and economy in poetic language not on the basis of an analogy with prose, but on the basis of the laws of poetic language" (ibid, p. 14). What would ensure the greatest economy would be algebraization or automation, something completely alien to poetry which is supposed to deliver us from the automatism of perception.

For Shklovsky, poetic language as such poses only difficulties (ibid., p. 26), profoundly extending the process of perception, opposing the principle of language's economy. The nature of poetic language indicates that it is, in a sense, a foreign language. It is confusing, bizarre, improper, creates communication barriers, and, as such, it is programmatically uneconomical, retarding – cultivating the principle of expenditure, not economy. From the point of view of economics, therefore, what Shklovsky considers to be appropriate for poetry and Markowski for modern literature – is the very same thing.

The economic categories used by the formalists are, above all, based on thinking in terms of colloquial language as a simple exchange which causes the word in circulation to be consumed, conventionalized, automatized, thus losing its autonomy, and, in the end, perceptual properties. This pessimistic diagnosis, which

leads Shklovsky to a vision of "extinct words" and a "language cemetery," allows nonetheless for a longed-for alternative, which is to be embodied in poetry, saving language from conventionalization and automatization of perception. If everyday communication is governed by exchange, with the currency of words in constant circulation, then poetry will be the point at which this cycle is interrupted and the currency falls out of circulation, regaining autonomy relative to communicative coercion.

Jean Baudrillard starts from a similar catastrophic vision of a great language garbage heap where all metaphors conventionalized in everyday communication end up. He formulates one of the most interesting descriptions of how the economics of poetry operates, combining the structural concept of the sign with the mechanisms of political economics (Baudrillard, 2006). Baudrillard makes semiotics the starting point in his argument (starting from de Saussure through Jakobson and Kristeva), showing the "subversive influence of poetic form on linguistics" (Baudrillard, 2007, p. 280) and the fact that semiotics is essentially heading towards "more or less subtle suppression of the radical character of poetic language" (ibid.), where poetic language, according to Baudrillard, is definitely subversive. As Michał Kłosiński notes, Baudrillard "throws a revolutionary torch at the spot where it causes the greatest havoc – on the pile of poetry, hence the chilling title of his dissertation: *The Poetic as the Extermination of Value*" (Kłosiński, 2015, p. 66). Poetic language in this approach resists the economic mechanisms to which, like the financial system, the system of language signs is subjected. Poetics should, therefore, be understood not in terms of economic exchange, but as a form of symbolic exchange. In this sense, Kłosiński concludes that for Baudrillard, "poetic language establishes an enclave of symbolic exchange, it defies the law of values in which the stake is the accumulation of meaning, without which it is impossible to play a speculative game with signs that are detached from reality. Baudrillard ... tries to save poetic language from the ruthless economic policy of the sign" (ibid., p. 69). He notes that poetic language will not be thus subjected to the laws of signs-to-things exchange, but will remain a unique gift offered by a text as part of symbolic exchange. The most mysterious here, however, is the vision of the "self-immolation" of the sign, described as "an attempt to be saved from the threatening inflation of meaning" (ibid, p. 73) finding its end in the complete conventionalization

of metaphors, a "wearing-out" of language, rather like an old coin remaining in circulation for too long. In the sphere of poetic language, signs are therefore not exchanged like money and goods, but like gifts, similarly to primitive societies. The purpose of this exchange is not profit but the negation of practical value which is fulfilled by the demonstrative destruction of what is given[10].

Baudrillard's theory, presenting poetry as a space for symbolic exchange, an alternative to the political economy of the sign, became one of the most important inspirations for the Italian theoretician of culture and philosopher, representative of the Italian Theory, Franco "Bifo" Berardi. His vision of the "economics of poetry" is based on belief in its critical and emancipatory potential. In such an approach, poetry would re-create itself as a territory subjected to the mechanisms – once again – of another, alternative economics, becoming the point of resistance against the oppression of various types of systems. Berardi develops his entire project on the economics of poetry around this premise. In his book *The Uprising. On Poetry and Finance*, he analyzes how capitalism, with its oppressive economic discourse, leads to the subordination and automation of language and the subjugation of affective potential. Berardi thus shows the subversive possibilities of poetry, accompanied by a unique project of uneconomicality and sensitivity, exhibiting critical and emancipatory properties at the same time.

One of the keywords of Berardi's theory is a kind of state of "bankruptcy," i.e. "insolvency," which he sees as being synonymous with becoming independent of the hierarchy of values and the list of priorities connected with the capitalist system and neoliberal conformism. According to Berardi, from a linguistic point of view, insolvency would set a possible escape route from the reduction of language to exchange (Berardi, 2012, pp. 16–17). By showing the social organism as one that entails a network of "techno-linguistic" automatism and describing processes that contribute to language automation, Berardi depicts poetry as "an excess of language" (ibid., p. 22), whose basic property is

---

[10] The theory of the anagram, being an alternative to structural poetics, and two laws described by Ferdinand de Saussure – the law of the word-theme and the coupling – are crucial for Baudrillard (Baudrillard, 2006, p. 320; Dziadek, 2001; Kłosiński, 2018, p. 78).

non-exchangeability, creating conditions for de-automatization (ibid., p. 20). Poetry creates new, alternative worlds that can form the basis for a community of thinking and understanding. It is shown as the vibration of a single voice, endowed with the possibility of resonating, and while it resonates, it can generate a space for the community.

In Berardi's approach, poetry transcends the boundaries set by ideological linguistic exchange, showing the possibility of a new codification of the relationship between the signifier and the signified, both in a nonrecurring, unique voice and through a nonrecurring, unique listener. Following Giorgio Agamben's reflection on the voice being the meeting point of body and meaning, Berardi states that poetry is "the voice of language" (ibid., pp. 20–21). In this perspective, the metaphor of insolvency reveals its extraordinary potential. The right to insolvency is not only a figure of social or civic resistance to the market entanglements of entities, a refusal to participate in repayment of a real and metaphorical financial debt. It is, above all, a refusal to subordinate living potential to the domination of a formalized, politically legitimized economic code allied with mechanisms of biopower. The capitalist, neoliberal, ideologized form of thinking and language is therefore not only a set of economic rules. It is also a variety of different internalized borders beyond which our imagination should not venture, and which are, at the same time, the framework of linguistic automatism – easy to subjugate and control.

As Berardi claims, we have to think of a different theory dealing with the perception of reality, one beyond the fixed ideological dictionary (ibid., p. 147). We should try to imagine the liberation of potential from the power of neoliberal forms of organization governing our reality experience, from automation ruling everyday life by means of linguistic form. For Berardi, all of this could be achieved by poetry – all that is poetic is a form of linguistic insolvency, enabling this gesture of resistance and rejection. This "insolvency" being the basic property of poetic language would mean, among other things, "the rejection of the economic code of capitalism" (ibid., p. 149) that keeps real life and social potential locked within form.

By not taking part in the exchange of things for words according to matching codes, poetry disarms the transparent relationships of consent, allowing us to embrace the undefined and

attempt to redefine the world. As such, it is a space for shaping the sensitivity of the community and a chance for a different perception of reality.

There is no doubt that Bernardi's concept of language de-automatization was borrowed from Shklovsky (Berardi, 2012, pp. 149–150) who, in a similar vein, recognized the potential of poetry as being related to "removing objects from the automatism of perception" (Shklovsky, p. 16). In turn, the very concept of verbal "insolvency" is closely related to Baudrillard's vision: symbolic exchange in which the rules of political economy are compromised. In Berardi's project, two theories of poetic language are intertwined, crucial from the point of view of the economics of poetry. They become the starting point of a political project where the social significance of the poem is at stake.

In each of the selected but only briefly illustrated theoretical approaches, the contradictory diagnoses are surprisingly harmonized within the perspective of the economics of poetry, based every time around the concept of poetic language being uneconomical, which is usually due to the interruption, or even suspension, of the circulation of words. The uneconomical nature of poetry becomes vital every time for recognizing its exceptional value which, at the same time, has nothing to do with situating it in a role resembling a priceless but useless jewel. In the case of each of the several approaches outlined here, all of them different but still engaged in a dynamic dialogue, poetry can be defined as a movement of resistance to conventionalized communication and automation of perception, a medium offering unique knowledge about reality, a possible space for language's emancipation from the rules of economics as well as communicative coercion. Poetry remains a space that unites the community of symbolic exchange and, finally, it is a projected area of subversion that helps liberate our experience of reality from neoliberal forms of organization and offers a different mode of perception.

Following Aristotle in diagnosing the unique properties of metaphor, Jacques Derrida once again takes us to an unattainable treasure at the end of the rainbow: "In nature, everyone has his nature. ... If the greatest thing by far is to be a master of metaphor, some have the gift of metaphor, know better than others how to perceive resemblances and uncover the truth of nature. A capacity not within our grasp" (Derrida, 2002, p. 304). Sounds like invaluable usefulness!

* * *

In this way, the economics of poetry seems to shed light on the cognitive potential of the poetic medium. It also seems to open the possibility of a new definition of specifically poetic knowledge about reality, urging us to rethink traditionally understood duality of form and content that isolates a text from its historical context, placing poetic work and the aesthetic impact of poetry outside the area of social practices. It might offer a chance to prevail over approaches in which form becomes a non-exchangeable residue, breaking its connection to reality, declining participation, involvement and responsibility for its shape – against the vision of the cognitive, critical, social and political uselessness and helplessness of poetry. Studies on the economics of poetry create the possibility of introducing a heteronomous poetic form in which poetry at the same time becomes a medium that deposits a specific type of knowledge; a place of articulation of individual and communal experience inextricably linked to the poetic form, invoking different images of the past, at the same time becoming a prism through which both the present and alternative projections of the future can be seen.

To conclude – a word about a colorful, almost rainbow-like bouquet of flowers. We find it in Banksy's mural depicting a masked man; he looks like an assailant, his pose that of someone preparing to throw a grenade or Molotov cocktail. The only difference is that instead of a potential tool of crime or destruction he is holding a bouquet of colorful flowers. Of course, this mention of *Flower Bomber*, as the mural is usually referred to, is not random. In an interview with Grzegorz Jankowicz, *Niewielki odwet na „prawdziwym" życiu* [Small Revenge on "Real" Life], Andrzej Sosnowski admits that Banksy's mural would be great for the cover of his volume of poetry. Is this, then, how the poet sees the role and functioning of poetry? He explains to his interlocutor: "Revolutions imagined in poetry can only be bloodless. The uselessness of the most interesting pieces has always seemed to me their naturally priceless side. Flowers, poetic tropes – they are almost like synonyms of awkwardness and ineffectiveness throughout the times" (Sosnowski, Jankowicz, 2010, p. 191). Sosnowski equates the uselessness and pricelessness of poetry which, being deprived of practical and exchange worth, cannot be exchanged or valuated, and that is exactly why it can

operate outside economic rules. Poetry's paradox lies in the fact that once degraded to being useless and ineffective, it does not become worthless – but priceless.

Translated by Marta Gorgula

References

Abrams, M.H. (2012). *Zwierciadło i lampa. Romantyczna teoria poezji a tradycja krytycznoliteracka* (Polish edition of *The Mirror and the Lamp: Romantic Theory and the Critical Tradition*, transl. M.B. Fedewicz). Gdańsk: słowo/ obraz terytoria.

Barańczak, S. (2007). *Wiersze zebrane* [Collected Poems]. Kraków: Wydawnictwo a5.

Barthes, R. (1999). *S/Z*, transl. M.P. Markowski, M. Gołębiowska. Warszawa: Wydawnictwo KR.

Baudrillard, J. (2007). *Wymiana symboliczna i śmierć* (Polish edition of *L'échange symbolique et la mort*, transl. S. Królak). Warszawa: Sic!

Berardi, F. (2012). *The Uprising. On Poetry and Finance.* Cambridge, MA: MIT Press.

Blumenberg, H. (2017). *Paradygmaty dla metaforologii* (Polish edition of *Paradigmen zu einer Metaphorologie*, transl. B. Baran). Warszawa: Wydawnictwo Aletheia.

Compagnon, A. (2004). *Literature, Theory and Common Sense*, transl. C. Cosman. Princeton: Princeton University Press.

Derrida, J. (2002). Biała mitologia. Metafora w tekście filozoficznym (Polish edition of *La mythologie blanche: (la métaphore dans le texte philosophique)*, transl. J. Margański). In: J. Derrida, *Marginesy filozofii* (Polish edition of *Marges de la philosophie*, transl. A. Dziadek, J. Margański, P. Pieniążek) (pp. 261–336). Warszawa: Wydawnictwo KR.

Dziadek, A. (2001). Anagramy Ferdinanda de Saussure'a – historia pewnej rewolucji [Ferdinand de Saussure's anagrams: The history of a revolution]. *Teksty Drugie*, No. 6.

Felski, R. (2016). *Literatura w użyciu* (Polish edition of *Uses of Literature*, transl. J. Borkowska et al.). Poznań: Wydawnictwo PSP.

Foucault, M. (2005). *Słowa i rzeczy. Archeologia nauk humanistycznych* (Polish edition of *Les mots et les choses: Une archéologie des sciences humaines*, transl. T. Komendant). Gdańsk: słowo/obraz terytoria.

Hörisch, J. (2000). *Heads or Tails: The Poetics of Money.* Detroit: Wayne State University Press.

Jakobson, R. (2006). Poetyka w świetle językoznawstwa (Polish edition of Linguistics and Poetics, transl. K. Pomorska). In: A. Burzyńska, M.P. Markowski (Eds.), *Teorie literatury XX wieku. Antologia* [Literary Theories of the 20th Century. An Anthology]. Kraków: Wydawnictwo Znak.

Jameson, F. (2009). Mapowanie poznawcze (Polish edition of Cognitive Mapping, transl. B. Kuźniarz). *Krytyka Polityczna*, No. 16/17, 232–240.

Kłosiński, M. (2015). *Ratunkiem jest tylko poezja. Baudrillard – Teoria – Literatura* [The Only Hope Is Poetry. Baudrillard – Theory – Literature]. Warszawa: Instytut Badań Literackich PAN.

Markowski, M.P. (2004). Reprezentacja i ekonomia [Representation and economics] (pp. 11–27). *Teksty Drugie*, No. 4.

Markowski, M.P. (2007). *Polska literatura nowoczesna. Leśmian, Schulz, Witkacy* [Polish Modern Literature: Leśmian, Schulz, Witkacy]. Kraków: Universitas.

Moretti, F. (2016). *Wykresy, mapy, drzewa. Abstrakcyjne modele na potrzeby historii literatury* (Polish edition of *Graphs, Maps, Trees: Abstract Models for a Literary History*, transl. T. Bilczewski, A. Kowalcze-Pawlik). Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Nietzsche, F. (1993). *O prawdzie i kłamstwie w pozamoralnym sensie* (Polish edition of *Über Wahrheit und Lüge im aussermoralischen Sinne*, transl. B. Baran). In: F. Nietzsche, *Pisma pozostałe 1862–1875* [Other Works 1862–1875]. Kraków: Inter Esse.

Nycz, R. (2001). *Literatura jako trop rzeczywistości. Poetyka epifanii w nowoczesnej literaturze polskiej* [Literature as a Trope of Reality. Poetics of Epiphany in Modern Polish Literature]. Kraków: Universitas.

Shell, M. (1978). *The Economy of Literature*. Baltimore: Johns Hopkins University Press.

Sieburth, R. (1995). Ezra Pound: Ekonomia poezji/ poezja ekonomii (Polish edition of In Pound We Trust: The Economy of Poetry/The Poetry of Economics]. *Literatura na Świecie*, No. 1–2.

Sosnowski, A., Jankowicz, G. (2010). Niewielki odwet na „prawdziwym" życiu [Small Revenge on 'Real' Life]. In: G. Jankowicz (Ed.), *Trop w trop. Rozmowy z Andrzejem Sosnowskim* [Trope after Trope. Conversations with Andrzej Sosnowski] (pp. 173–192). Wrocław: Biuro Literackie.

Szkłowski, W.B. (1986). Sztuka jako chwyt (Polish edition of Shklovsky, V.B., *Iskusstvo kak priyom*, transl. R. Łużny). In: S. Skwarczyńska (Ed.), *Teoria badań literackich za granicą. Antologia* [Theory of Literary Studies in Other Countries. An Anthology], vol. 2 (pp. 10–28). Kraków: Wydawnictwo Literackie.

Tomczok, P. (2018). *Literacki kapitalizm. Obrazy abstrakcji ekonomicznych w literaturze polskiej II połowy XIX wieku* [Literary Capitalism. Images of Economic Abstractions in Polish Literature of the Second Half of the 19th Century]. Katowice: Wydawnictwo Uniwersytetu Śląskiego.

Ulicka, D. (2006). O funkcji poznawczej literatury i wiedzy o literaturze (tezy do przyszłej antropologii literaturoznawstwa) [On the cognitive function of literature and literary knowledge (theses for a future anthropology of literary studies)]. In: W. Bolecki, E. Dąbrowska (Eds.), *Literatura i wiedza* [Literature and Knowledge] (pp. 9–32). Warszawa: Wydawnictwo IBL PAN.

Woodmansee, M., Osteen, M. (Eds.) (1999). *The New Economic Criticism. Studies at the Intersection of Literature and Economics*. London: Routledge.

Paweł Tomczok
ORCID 0000-0003-3618-4844
University of Silesia in Katowice

Truth and Falsehood of the Mirror: Subjectivity – Reflection – Practice

Abstract

The author discusses different uses of the mirror metaphor in philosophy, literature and aesthetics. On one hand, a mirror serves to talk about human empowerment and inner life, while on the other becoming one of the most important metaphors of truth, of proper representation of reality. The final part of the paper considers the practical location of mirrors – and the practice of mirroring as the basis for other philosophical and existential metaphors. The author treats the mirror not only as a metaphor but also as an artefact with its own history: the history of an object but also social and economic history. By taking into account this physical history of mirrors, we can observe the changing foundation of epistemological and aesthetic metaphors.

Keywords: mirrors in culture, dialectics, metaphorology

Mirror, reflection, mirroring. Both the process of reflection and the artefact itself create an imaginational resource of countless philosophical models of truth and falsehood. A mirror-reflected image suggests it is possible to produce a faithful representation of reality, but also – almost at the same time – that this image can be inaccurate, distorted, blurred, too dark or too bright. Thus the mirror defines the purpose of cognition as well as simultaneously indicating the possibility of different obstacles to achieving that purpose.

Of course a mirror is not the only metaphor of truth and falsehood; another equally important example would be the image of light and darkness, organized around the metaphor of, for instance, a lamp, which is the mirror's great rival in both science and the arts (see Abrams, 2003). Quoting these two examples enables us to notice the real, physical basis of both metaphors,

which changed their meaning with the development of technology – with the improvement and popularization of artificial reflections as well as artificial light sources (Schivelbusch, 2004). If we also remember that both these technologies are often combined, for instance in mirror rooms where mirrors multiply the light by dispersing it, it will turn out that the two metaphors do not have to constitute alternatives and rivals at all, but are open to mutual translation. Indeed, light is a prerequisite for a mirror-reflected image, while mirrors themselves also contribute to brightening up various rooms by helping light disperse.

The present paper will discuss three fundamental meanings of the mirror in different discourses. The first two refer mainly to cognition and experience, to learning and aesthetics. A stable mirror standing in one place has become one of the main metaphors through which we speak of human subjectivity, and also of learning about the world. The experience of seeing our own mirror image is one of the pivotal moments in the development of human subjectivity – while the story of Narcissus gazing at his reflection has been an influence all the way to 20th-century theories of primary and secondary narcissism characterizing the constitution of individual consciousness, and in the theory of collective narcissism describing how a collective subject views itself in the mirror. Thus, the mirror is an artefact producing and disseminating different techniques of reflection – not only of the self but also social entities, networks, which reproduce themselves through a series of reflections. Therefore, the first part of the paper will discuss different subjective mirrors.

Part two will present different forms of mirroring – above all in art and philosophy. Mirroring is one of the most important aesthetic models: According to many theorists, art always reflects a reality, while others believe art should also mirror reality, because otherwise it does not fulfill its task. Thus, mirroring defines models of descriptive as well as normative aesthetics. A similar duality can be identified in metaphysical thinking about human cognition: Reflecting on the human mind leads fluidly to seeking adequate ways of mirroring the world. The mirror metaphor, therefore, stretches between the obvious situation of a direct reflection in a mirror and the realization that this reflection is not simple and obvious at all, as it encounters many possibilities of distortion, falsification, resulting in a reflection that is inaccurate, untrue, that has to be amended, corrected, changed.

The first two sections will focus on the subjective, theoretical and aesthetic meaning of the mirror. The model here will be mainly mirrors situated in one place, stable mirrors as well as small mirrors in which we see our own face – look at ourselves. In part three of the paper, I would like to analyze the less intuitive although also important and widespread meaning of the mirror, i.e. its practical application. Mirrors in everyday use are not only for looking at ourselves in detail; they also expand the human field of vision. They enable us to see what is behind us, to see things we cannot see when we are looking forward. Examples of such practical mirrors include car mirrors enabling drivers to see vehicles behind them. Thus, these are mirrors providing knowledge on changing reality, and at the same time causing the multiplication of different reflections in reality.

In all the sections of the present paper, the mirror and the reflection process are considered in a dual sense. At the forefront are metaphors – but we must not forget that those metaphors have a physical foundation that also underwent change. The metaphorological history of the mirror is the history of both the idea and the artefact. New possibilities of reflection, increasingly accurate and greater, move the process of mirroring away from its natural, fluid character toward fixing the image. This history of the technology of reflection is a social and economic history as well: It is also the history of the price of mirrors, monopolies tied to their manufacturing, the history of tradesmen who knew the secrets of making them, and, finally, the history of the democratization of mirrors, when everyone can view their own image at any time.

This complicated, network-like history of metaphors proceeds toward treating them much more broadly than in various linguistic and cognitive approaches. It is precisely examples like the mirror (and also light) that suggest we are dealing with more than metaphors, even than absolute metaphors to which Hans Blumenberg ascribed the special role of representing the entirety of philosophical ideas like the idea of the world or God, history, and humans in relation to them (Blumenberg, 2013, 2014). Those metaphors and artefacts serve as mediators taking part in negotiating different ways of understanding reality. Thus, these kinds of objects are not just metaphors serving to build utterances, but they also have the function of a technology for representing reality. They enable us to reduce complex reality by creating nodal points in the knowledge network.

1. The mirror: histories of subjectivity

In his *Dzieje zwierciadła* [*The History of the Mirror*] Mieczysław Wallis notes that "the invention of the mirror dates back to times immemorial, it preceded the invention of writing" (Wallis, 1974, p. 11). However, the mirror is an invention while also being something natural: Before the first polished metals and stones in which you could see your reflection appeared, people could see mirror images in water. The mirror as an artefact immediately puts humans in competition with nature, in a contest for creating a better, more accurate and more lasting mirroring of reality, particularly humans themselves. This suspension between nature on one hand and technology and artificiality on the other makes the mirror similar to another phenomenon that is both technical and natural, i.e. fire. Insofar as fire has the myth of its theft, a myth that became the foundation of many emancipative historiosophies, the mirror seems not to have played such a significant role. The Narcissus myth refers us more to what is private and intimate, while mirrors themselves – despite their important social role – have interfered more strongly with the individual sphere.

In the historiosophy of modernity as well, fire seems to play a more important role: It was a new way of controlling fire that launched the industrial revolution, while the burning of coal (among other things) enabled Europe to gain an advantage over the rest of the world in the 19th century and build a hegemony lasting many years. In this Promethean model, it is easy to forget the mirror. Nevertheless, starting from the 17th century we can find someone in practically every period who claimed it was the Baroque, the Enlightenment, or the 19th century that constituted the age of the mirror. This apparent discrepancy stems from the democratization of the mirror, which progressed with the process of modernity. Back in the 17th century, during the Venetian monopoly on mirror manufacturing, mirrors were extremely rare and precious goods, accessible mainly to the aristocracy (Melchior-Bonnet, 2007). It was not until new production technologies were developed that mirrors could be made in greater numbers, placed in large spaces, and their price could go down. Thanks to these complicated processes, mirrors passed from courtly society, where they had enabled people to train their person in proper presentation among high society,

to the world of burgher intimacy. Whereas the Promethean myth was responsible for a comprehensive story about the progress of civilization, the mirror became a medium for shaping modern subjectivity which, through different reflections, has had to undergo continuous self-observation to this very day.

Already the story of Narcissus found in *Metamorphoses* presents a situation of looking at the water reflection of one's own face – recognizing oneself, one's face, opens the way to thinking about identity, self-knowledge as well as different affects that a subject can direct toward himself or herself. Thus, a mirror also becomes the start of self-reflection – let us note, however, that the source of this self-reflection is not in the subject, in some supposed inner life, but in a complicated relationship with objects, with surfaces capable of producing reflections. The Narcissus myth involves a dual process: In the foreground we have the protagonist's mirror duplication. Therefore, the mirror creates a strange relationship: not with others – with oneself, but as a duplicated being, seen in reflection, and at the same time felt from the inside. It needs remembering, though, that to see himself, Narcissus had to isolate himself from the group. This entangles us in the complicated dialectic of individualization and duplication, a dialectic in which the subject abandons relationships with others in order to meet the Other as his or her own reflection.

The individual mirror mechanism has been processed in different ways in psychological concepts: in Sigmund Freud's notion of narcissism, Jacques Lacan's mirror stage, and in recent decades in the metaphor of mirror neurons (Keysers, 2017). All these theories assume we process our relationship with ourselves through external reflections; they make a shift in the strict separation between the individual and everything outside, especially other subjects, and also animals. Passing the mirror test was supposed to be one of the more important stages in becoming human. Similarly, the concept of mirror neurons locates the uniqueness of humans in their capacity for empathy and the development of special social competence enabling them to form numerous and complex groups capable of knowledge transfer and cooperation (Tomasello, 2017).

The theory of narcissism corresponded to different models of individualism and individuality. A broader version of this approach could be seen in the theory of collective narcissism outlined

by Erich Fromm in *The Anatomy of Human Destructiveness*. In group narcissism, the object of adoration is the community looking at itself in various symbolic images of itself; this collective fascination with one's group identity results in fanaticism, manifesting itself especially strongly in sensitivity to collective symbols. They are precisely the mirrors in which the community wants to see itself. Any violation or criticism of such symbols is perceived as an attack on the identity of the group constituted around those images.

The mirror image of the subject proceeded toward becoming more permanent. From a reflection in water, whose surface is the easiest to disturb, we moved on to artificial mirrors which, however, did not present an accurate image for quite a long time, so it was no wonder that the mirror became one of the most important symbols of inaccurate cognition, contrasted with directly looking at one another face to face. Photography seems to be the most important way of fixing a mirror image, and in its successive versions it enabled images to be captured more and more accurately and quickly. Obviously, digital forms of such stabilization of mirrored subjectivity include selfies, as images that you can see and capture in an instant and then disseminate effortlessly. In this case, we manage to eliminate the difference between a mirror reflection, always fleeting and transitory, and an image, which usually involves a temporal delay and other forms of indirectness. This model also defines new boundaries of mirror self-control accomplished thanks to the possibility of continually looking at oneself in captured images. An image of oneself, which is often modified and improved, becomes a template the subject tries to fit into.

Different meanings of the subjective mirror are analyzed by Heinrich Kleist in his essay *On the Marionette Theater*. This short text is based on the difference between a conscious being and a puppet or marionette, as it is structures of the human body deprived of consciousness that supposedly have the greatest charm. In the narrator's conversations with the mysterious Herr C., there appear three examples of confrontation between a conscious being and non-conscious bodies: the first one is the marionettes from the puppet theater, the second example is that of a statue that a graceful young man tries to imitate, while the third concerns an incredible swordfight with a master-swordsman bear. Example number two features a mirror, so let me quote it:

…we had just recently seen the statue in Paris of the youth removing a thorn from his foot; the statue is well-known and models of it can be found in most German collections. A look cast into a large mirror, at the moment when he was placing his foot on a stool to dry it, reminded him of the statue; he smiled and told me of the discovery he had made. In fact, I had, just at that very moment, made the same discovery; however, whether it were to test the sureness of the grace that possessed him, or whether it were to cure him a little of his vanity, I laughed and replied that he was probably seeing ghosts. He started to blush and raised his foot a second time to convince me; but the attempt, as could easily have been foreseen, was unsuccessful. He raised his foot in confusion a third and a fourth time, he raised it probably another ten times: all to no avail! He was incapable of repeating the same movement. What am I saying? The movements he made had such a comical element to them that I had difficulty keeping from bursting out laughing.

From that time, from that very moment on, an indescribable change came over the young man. He started to spend his days standing in front of the mirror; and as he did so one attractive feature after another deserted him. An invisible and unaccountable power seemed, like an iron net, to lay itself over his gestures and facial expressions, and after a year had passed, there was no trace left to find in him of the loveliness that the eyes of the people, who otherwise surrounded him, had delighted in[1] (Kleist, 2000).

The equal participants in this quite complex scene are two young men and two objects: an ancient statue and a mirror. Each of these actants influences the others. First of all there is the statue, treated not as a reflection of reality but rather as an ideal, a model that requires representation. One of the young men successfully imitates the ancient statue – in rather a complicated arrangement: He notices the correspondence in a mirror; also his friend does not see the reproduction of the statue's stance in reality but only as a reflection. But this fleeting mirroring is gone immediately; it is easy to question but impossible to repeat.

Attempting to repeat it results in the loss of gracefulness. Completely different notions appear in its place. The young man's body becomes imprisoned – although there is no question of any

---

[1] English text from the translation by Kevin J M Keane, retrieved from https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=7&ved=2ahUKEwi8jbqXoqrjAhVhl4sKHdX8AnkQFjAGegQIBxAC&url=http%3A%2F%2Fwww.kevinjmkeane.com%2Fwp-content%2Fuploads%2F2015%2F07%2FKleist-On-the-Marionette-Theatre.-July–2015.pdf&usg=AOvVaw26Ps8BWdl4oDxC7OuQrRHP

physical imprisonment, the metaphors used suggest something heavy and restrictive: An invisible, incomprehensible power appears as an iron net. This set of metaphors is difficult to keep separate from economics – an iron cage, a structure hard as steel, and the invisible hand of the market that determines people's economic actions. The young man in front of the mirror thus loses his charm and becomes entangled in the net of bourgeois life, a life without charm, a life tied to daily work and duties. In this case the easily accessible, private mirror turns out to be not so much a tool of the aesthetic narcissism of a romantic, beautiful soul as a mediator involving the individual in economic practice: The mirror, contrary to the ancient statue, is a commodity you have to pay for, and to pay for it you have to become implicated in a network of economic relations.

2. Mirror – mirroring of the world – art and science

The mirror model is one of the main tools of dialectic philosophy (Holz, 2005; Schickel, 2012). Leibniz's monadology treats an individual monad as a living reflection of the entirety of the world, and the relationship between what is mirrored and the mirroring itself opens up a whole complex of problems linked to the temporal and spatial aspect of representation. Contemporary philosophy in some of its realizations has moved from dialectic relationships to normative demands – then, mirroring stops being a dialectic relationship and becomes a cognitive and aesthetic norm. In such a case, science and art are given the task of correctly mirroring reality, for example anything in reality that is typical or progressive.

Note that in such a case, reification of the metaphor takes place – the reflection relationship is treated as virtually an automatic mechanism, to the exclusion of the mirror, or even that mirror moment, understood as the place of a separate artefact, the spectacle's mediator. In this theory, physical matter became primordial, while the social world was only supposed to be its reflection – this turned reflecting into a relationship of temporal succession and causality. Scientific learning about the social world and culture was meant to be limited to identifying relationships with physical entities. In aesthetics, too, the vision of the mirror

wandering across a courtyard to reflect all kinds of objects and phenomena can be reduced to the demand of simple realism, which is meant to be the criterion for evaluating individual works judged by their degree of progressiveness.

Where does this frequent reification of the metaphor come from? It occurs when the metaphor wears down, when you can no longer sense the moment of comparison, i.e. when the complicated play of different compared objects is replaced with the illusion of the directness of a relationship, e.g. between language and reality. Meanwhile, living philosophy makes bold use of different objects that continue to have the metaphorical power of opening up to things that are new and unexpected, thus often drawing upon various technical and media-related innovations. This is the case, for instance, in the well-known passage from *The German Ideology*:

> If in all ideology men and their circumstances appear upside-down as in a *camera obscura*, this phenomenon arises just as much from their historical life-process as the inversion of objects on the retina does from their physical life-process[2] (Marks, Engels 1961, p. 27).

The definition of ideology here has nothing to do with scientific exactitude, with serious defining of notions. Instead of logical operations, Marx builds a network of mediator objects: a *camera obscura*, which he must have known from experience but also from various diagrams explaining how it works, as well as the knowledge of optics and physiology of the time, also mediated through drawings of the eye. Therefore, this sentence is backed not only by a simple simile, but also by the institution of science from the first half of the 19th century, the latest media, and also the rules for copying the reality studied by science into textbooks and scientific works. It is something Bruno Latour

---

[2] English text from Karl Marx & Frederick Engels, *The German Ideology: Part One with selections from Parts Two and Three, together with Marx's "Introduction to a Critique of Political Economy"*, edited and with introduction by C. J. Arthur, 2004, New York: International Publishers, retrieved from https://books.google.pl/books?id=DujYWG8TPMMC&pg=PA47&lpg=PA47&dq=%22if+in+all+ideology+men%22&source=bl&ots=j_VcXqrvUZ&sig=ACfU3U226-PEZWuEdCSVpzckxa_SuHkQFg&hl=en&sa=X&ved=2ahUKEwi22Prh4qrjAhVKtIsKHdI1CpIQ6AEwA3oECAgQAQ#v=onepage&q=%22if%20in%20all%20ideology%20men%22&f=false

called procedures of inscription – inscribing reality into new means of representation. Dialectic mirroring of reality, therefore, cannot be considered as an abstract process but has to take into account a network of different artefacts, mediators taking part in creating the representation.

The mirror also became the main metaphor in speculative philosophy, enabling us to speak of indirectness (Kuhn, 2014, p. 385). The structure of a mirror reflection involves a number of moments and differences that enable the content of speculative thinking to be expressed. This speculative status already belongs to the dual meaning of the mirror – on one hand being a physical object built from a specific material, while on the other being able to be a mirror only when it reflects something else – which makes it an object that actually cannot be seen, because other objects are always visible in it, reflected in the mirror surface. Being a mirror, therefore, is inseparably linked to the fact that we always see in it something other than the mirror itself. This structure has enormous speculative potential, because it enables us to think about objects that are visible things but at the same time are not so much the object that we are looking at, but rather they enable us to see something else, and, continuing – they make possible the existence of a reflection as well as enabling thinking about the difference between the thing reflected and the reflection. This means there is speculative potential in the duplication of objects itself, and in the problem of things that are visible and invisible at the same time, catching our eye in order to allow us to see something else.

3. Practical reflection

Both in the philosophy of the subject and in metaphysics and art, the mirror metaphor proceeded toward stability and immobilization. The mirror the subject faced, just like the reflection of the entirety of the world, was supposed to offer the possibility of aesthetic and theoretical cognition. The practical use of mirrors remained in the background. Meanwhile, mirrors do not so much serve the purpose of looking at ourselves or contemplating mirrored images as they expand our momentary perceptual capacity, enabling us to see what is "at the back of our head." In this context, Umberto Eco speaks of mirrors as being prosthetics, instruments expanding the scope of operation of a specific

organ, in this case the eyes (Eco, 2012, p. 18), allowing us to look behind us without turning our head.

Eco analyzes the example of an arrangement of mirrors thanks to which you can see what is going on behind you – there are movie examples of seeing someone entering a room and being noticed by another person. Probably the most interesting case, though, could be the use of mirrors in cars: Mirrors assist those traveling by car, where such travel in itself expands human capacity. Glancing in the mirrors is a routine action enabling the driver to gain information about other road users. Therefore, it is an important channel of communication with the surroundings. In this case the reflected image does not serve any aesthetic or theoretical contemplation, but is combined with the movement of body and vehicle. Mirroring serves to provide information important only at a given moment, connected with the current situation.

In theoretical deliberations on reflections, one crucial category was isomorphism as well as the structural correspondence between the thing represented and the representation. This shared form, however, assumed the hylomorphism of traditional metaphysics. But the mirror requires a different metaphysics – interesting inspiration in this respect is provided by Tim Ingold's project and his theory of medium, substances and surfaces. Above all, mirrors are special surfaces in which other surfaces are reflected – this applies to the water mirror as well as artefacts. It appears that mirrors create images without matter, but it may be more apt to state that these are images of surfaces without substance. It is hard to say that mirrors have the capacity to capture form and create some kind of isomorphism, especially in a metaphysical sense. It would be much more intuitive to speak of mirroring the surface of things. But what is this surface? In Ingold's concept (2018) the surface is the boundary between substance and medium – meaning that what is duplicated is a thing's surface boundary and not its internal structure. A relationship between different surfaces is formed – they are reflected in one another, creating a network of duplicates and appearances of spatial depth, flickering images multiplying the observed world.

Conclusion

Mirroring has functioned as the most important characterization of truth – both the truth of science and the truth of art.

The metaphor of mirroring, however, quickly got entangled in linguistic representation – it was words and sentences, notions and judgments that were meant to reflect reality. The dialectic of mirroring has gone in the same direction, treating the mirror as a principle of philosophical speculation. But the mirror can be treated completely differently – more physically and more network-based; then, it gains its own history – and it is precisely this history that always has to be somewhere in the background of the metaphor's interpretation. But the mirror, or actually the multiplicity of different mirrors and, more broadly, mechanisms by which reality is reflected in artefacts, is an important element of creating and representing reality. Mirrors become historical mediators that take part in a complicated network made up of people and things.

The dependence of theoretical thinking on metaphors, a popular notion in the second half of the 20th century, only partly fulfills the task of network-based thinking about mirroring. The proposals of Hans Blumenberg and Jacques Derrida (2002) emphasized the permanent presence of metaphors in philosophical texts, especially wherever the meaning of the most important categories needed to be explained. The concept of absolute metaphors assumed there is a limited set of images that mediate in understanding ideas such as the world, the soul, or history. I would not want to question the achievements of these philosophers, but it seems to me that their analyses stopped at too early a stage. Metaphors in the analyses of Blumenberg and Derrida are caged in the realm of text; they are words rather than things. Meanwhile, many of the absolute metaphors have very concrete equivalents in the world of life, and their meaning is not exhausted in schematic images. Most of these objects have their history, they undergo different changes – one such example being the mirror, changing its symbolic meaning but also its material structure.

Thus, the mirror sits at the intersection of different discourses. The mirror is a node in a complicated network – a node combining theory and practice, science and aesthetics, technology and speculation. Thanks to this capacity for uniting seemingly distant realms, an analysis of the mirror could be a paradigm for network and complex analysis in which the inhomogeneous components of the world find a voice.

This attempt at outlining different meanings of the mirror also shows how complicated the discussion must be about

abstract categories that we try to express with the help of the image of the mirror. The issue of truth and falsehood is one of those issues that can appear to be simple, clear and obvious, but in fact require the engagement of all kinds of actants: human and non-human, and also the presentation of the processes in which, thanks to the involvement of different phenomena and technologies, all that is human is shaped.

References

Abrams, M.H. (2003). *Zwierciadło i lampa. Romantyczna teoria poezji a tradycja krytycznoliteracka* (Polish edition of *The Mirror and the Lamp: Romantic Theory and the Critical Tradition*, transl. M.B. Fedewicz). Gdańsk: słowo/obraz terytoria.

Blumenberg, H. (2013). Światło jako metafora prawdy. Przedpole filozoficznego kształtowania pojęć (Polish edition of Licht als Metapher der Wahrheit. Im Vorfeld der philosophischen Begriffsbildung, transl. Z. Zwoliński). *Kronos*, No. 2, 32–59.

Blumenberg, H. (2017). *Paradygmaty dla metaforologii* (Polish edition of *Paradigmen zu einer Metaphorologie*, transl. B. Baran). Warszawa: Aletheia.

Derrida, J. (2002). *Biała mitologia. Metafora w tekście filozoficznym* (Polish edition of *La mythologie blanche: (la métaphore dans le texte philosophique*)). In: *Marginesy filozofii* (Polish edition of *Marges de la philosophie*, transl. A. Dziadek, J. Margański, P. Pieniążek) (pp. 261–336). Warszawa: KR.

Eco, U. (2012). *Po drugiej stronie lustra i inne eseje. Znak, reprezentacja, iluzja, obraz* (Polish edition of *Sugli specchi e altri saggi*, transl. J. Wajs). Warszawa: wab.

Fromm, E. (2014). *Anatomia ludzkiej destrukcyjności* (Polish edition of *The Anatomy of Human Destructiveness*, transl. J. Karłowski). Poznań: Rebis.

Holz, H.H. (2005). *Weltentwurf und Reflexion. Versuch einer Grundlegung der Dialektik*. Stuttgart-Weimar: J.M. Metzler.

Ingold, T. (2018). *Splatać otwarty świat* [Binding an Open World], selection by E. Klekot. Kraków: Instytut Architektury.

Keysers, Ch. (2017). *Empatia. Jak odkrycie neuronów lustrzanych zmienia nasze rozumienie ludzkiej natury* (Polish translation of *The Empathic Brain: How the Discovery of Mirror Neurons Changes Our Understanding of Human Nature*, transl. Ł. Kwiatek). Kraków: Copernicus Center Press.

Kleist, H. (2000). O teatrze marionetek (Polish edition of Über das Marionettentheater, transl. J. Ekier). In: T. Namowicz (Ed.), *Pisma teoretyczne niemieckich romantyków* [Theoretical Writings of German Romantics] (pp. 573–581). Wrocław: Ossolineum.

Kuhn, K. (2014). Spiegel. In: R. Konersmann (Ed.), *Wörterbuch der philosophischen Metaphern* (pp. 380–392). Darmstadt: WBG.

Latour, B. (2013). *Nadzieja Pandory. Eseje o rzeczywistości w studiach nad nauką* (Polish edition of *Pandora's Hope: Essays on the Reality of Science Studies*), K. Abriszewski (Ed.). Toruń: Wydawnictwo UMK.

Marks, K., Engels, F. (1961). *Ideologia niemiecka* (Polish edition of *Die deutsche Ideologie*). In: *Dzieła* [Works], vol. 3, I. Strumińska (Ed.). Warszawa: Książka i Wiedza.

Melchior-Bonnet, S. (2007). *Narzędzie magii. Historia luster i zwierciadeł* (Polish edition of *The Mirror: A History*, transl. B. Walicka). Warszawa: bellona.

Schickel, J. (2012). *Der Logos des Spiegels. Struktur und Sinn spekulativen Metapher.* Bielefeld: transcript.

Schivelbusch, W. (2004). *Lichtblicke. Zur Geschichte der künstlichen Helligkeit im 19. Jahrhundert.* München-Wien: Fischer.

Tomasello, M. (2019). *Becoming Human. A Theory of Ontogeny.* Cambridge, MA: The Belkamp Press of Harvard University.

Wallis, M. (1973). *Dzieje zwierciadła i jego rola w różnych dziedzinach kultury* [The History of the Mirror and Its Role in Different Areas of Culture]. Warszawa: Wydawnictwa Artystyczne i Filmowe.

Michał Rogalski
ORCID 0000-0001-8225-7026
»Artes Liberales« Academy

The Disengaged Researcher as a Type: Truth and Probability in Studies on Religious Thought

Abstract

The paper discusses the issue of the intellectual historian's attitude to the subject of his research, especially if it takes the form of an intellectual conflict or controversy. It is a critical analysis of the possibility of assuming an objective attitude towards one's own research: without any valuation. Based on Aristotle's remarks on rhetoric, the author shows that the presentation of research results in writing is linked to taking a position towards them and giving them a place and meaning in the context of broader reasoning, which is synonymous with the operation of evaluation. The empirical context of the presented argument is provided by a self-analytical case study: the author's research on Catholic modernism in Poland. Reflecting on it, the author points to another source of the impossibility of objectivity, hidden in the hermeneutic method of research. Gadamer's postulate of writing history again and again for the sake of the changing historical context leads to a strong emphasis of the worldview of researchers themselves in their studies.

Keywords: truth, probability, intellectual controversy, history of ideas, Catholic modernism

> A modern philosopher who has never experienced the feeling of being a charlatan is such a shallow mind that his work is probably not worth reading (Kołakowski, 1988, p. 1).

Studying a given historical and philosophical issue seems not to be connected with the author's personal engagement in the research problem[1]. For historians of philosophy the position of

---

[1] The project was financed from National Science Center funding granted on the basis of decision DEC–2013/11/N/HS1/04769.

a disengaged, objective, unbiased researcher is convenient in many ways. It enables them to position themselves at a distance from the problems being analyzed, and to lend their research universal, timeless significance. It is a great pleasure for the writer of a book or paper to study something in a way no one ever has before, to describe a phenomenon so exhaustively that the findings remain valid regardless of the changing sociocultural context.

This approach to the tasks of the history of philosophy or history of ideas, or finally intellectual history, is understandable for many reasons. It has a lot in common not only with the Horatian *non omnis moriar*, but also with the esteemed tradition of 19th-century German historiography which, having discovered modern tools for studying sources for the humanities, concluded that its own results had been proved and were immune to the passage of time. In modern-day reflection, the myth of the objective historian has been critiqued in so many ways as to make them impossible to outline here (White, 1973; Koselleck 1985 [1979], pp. 73–155 White, 1987; Ricoeur, 2004 [2000], pp. 293–342;). Therefore I would like to consider not the whole issue, but one selected aspect that is especially important from the point of view of an intellectual historian. That aspect is making value judgments about the sides of the given historical conflict that is being described.

If an intellectual historian decides to outline an issue as an intellectual controversy, thus abandoning a biographical approach (describing different attitudes) and an encyclopedic one (a systematic description of the issue), he gives himself the task of reconstructing that controversy. Regardless of whether we are considering a clash of two opposite options or an entire palette of stances appearing in a debate, recreating a controversy requires us to show their mutual relationships and the argumentation networks forming as a result of their coexistence. In such an approach, every argument has its own inner logical structure that is often polemical toward some counterargument or else itself triggers another counterargument.

The researcher's task is not only to reconstruct individual steps in the reasoning, but also to judge them: indicate their cohesion or lack thereof, assess the explicatory power of the argumentation as a whole or its individual parts. Can this be done from an impartial position? And if so, how do you describe that impartiality and operationalize it? The metatheoretical position that is essential

for analyzing a given debate conducted on a theoretical level should no doubt possess a set of characteristics enabling personal engagement in the debate to be avoided. Or maybe it is just the opposite? Maybe it is impossible to assume any "meta" position without also being on a specific side of the dispute in question?

I would like to answer these questions by starting from my own experience with research on a specific issue from the history of Roman Catholic thought: the debate on Catholic modernism in the last decade of the 19th century and in the early 20th century (Rogalski, 2018). The reconstruction of an intellectual dispute appears to be a valuable tool for a historian, enabling the scholar to present a historical situation in its dynamic character, avoiding petrification in separate stances taken out of context. However, when this kind of description is used, questions inevitably need to be asked about the truth or probability of hypotheses, arguments, or entire argumentation strategies. Neither the category of truth nor the category of probability can be eliminated when using value judgments. How, then, are we to work with them?

Philosophical research as a form of rhetoric

One strategy is to strive for maximum objectivity. Even if the question of complete objectivity has to be seen as complex and hard to resolve unequivocally, you can try to perform such research steps that enable you intentionally to avoid getting involved in the dispute. If, for example, different sides in a controversy ask questions about the future of Catholicism and its relationship with changing contemporary culture, the researcher of course may choose not to give his or her own answer to such a question.

This is easier if the dispute occurred a century earlier; then, even having and giving one's own reply has no direct link to the dilemmas of the parties to that dispute. With some simplification, we could call those parties the modernist and antimodernist side, even though modernism is a set of attitudes within a similar general directive of thinking, without immediate relationships between them and impossible to simplify to a single model variant (Arnold, 2007, pp. 11–16).

Continuing, the researcher can also avoid prematurely judging the reconstructed stances in the dispute by trying to show how

the individual actors used notions and what rules governed their inferences. Such an approach enables the inference mechanisms to be reconstructed together with the notional images or worldviews on which they are based. Then, the dispute becomes a confrontation of different sensitivities, different ways of presenting the world or different ways of acting in the world. The researcher does not have to take anyone's side; the point, rather, would be to determine the constructional mechanisms governing the debate's development. Readers of this kind of description will thus follow the development of events, looking at clashing models of explaining the world, and it will be their decision whether to take anyone's side or – if this is possible – to refuse to favor any side.

This way of applying research methods and building narratives can be called a r h e t o r i c a l s t r a t e g y. The adjective "rhetorical" should be understood in the Aristotelian way. In his *Rhetoric*, the Greek philosopher defined the concept as follows:

> That rhetoric, therefore, does not belong to a single defined genus of subject but is like dialectic and that it is useful is clear – and that its function is not to persuade but to see the available means of persuasion in each case, as is true also in all the other arts (Aristotle, 1991, p. 35 [1355a]).

Researchers adopting a rhetorical strategy search for "available means of persuasion in each case." They analyze the spectrum of possible approaches to a given issue and place the positions assumed in the dispute within that spectrum. It is hard to accuse such an approach of lacking in objectivity. Objectivity is achieved by presenting all the possibilities, or a specific number of possibilities that are key within a given spectrum. This protects us from the necessity to assign veracity to any stance. Every position in a dispute is simply the realization of one of the intellectual possibilities whose construction the researcher is describing.

Does such inner theoretical discipline invalidate the question asked at the outset? Aristotle's very apt remark in his *Rhetoric* shows it does not. The philosopher writes:

> … for it belongs to the same capacity both to see the true and [to see] what resembles the true, and at the same time humans have a natural disposition for the true and to a large extent hit on the truth; thus an ability to aim commonly held opinions is a characteristic of one who also has a similar ability to regard the truth (Aristotle, 1991, p. 33 [1355a]).

Of course the remark that people have a natural instinct for striving for the truth can, for many reasons, discourage contemporary readers, who have moved away from those seemingly natural Aristotelian categories. However, even if we were to consider this one statement to be false, in my opinion that does not weaken the aptness of the belief in the same "faculty" of humans searching for truth and for probability. Presenting this observation using more contemporary terms, I would say that considering something to be true and considering something to be probable are two structurally identical actions. Their similarity means that, as far as the problem discussed here is concerned, switching the category of truth with the category of probability does not affect the existence of the issues outlined at the beginning. This brings us to the need to understand what position the researcher occupies in relation to the issues being studied. Regardless of what we call the value the researcher gives to the stances in the dispute under investigation, the problem of making value judgments is still there.

A historian of thought cannot avoid it, not even by means of maximum formalization. By formalization I mean using an analytical approach as consistently as possible, avoiding ambiguous notions and giving unequivocal definitions for any terms used, to enable the reasoning to be conducted with precision. Formalization is also linked to a tendency to replace categories of complex cultural origins or significant ethical implications with those that are devoid of such implications. For example, in such an approach, instead of assessing the "rightness" or "worth" of arguments, the researcher focuses more on their logical coherence or the simplicity of their semantic structure.

Formalization does not solve anything, however. The study process is not limited to thinking alone, but is also connected with communicating the results thereof. Every researcher has to assume this dual role: thinker and describer. In the case of studies on the history of thinking, we are dealing not only with a rhetorical philosopher whose activity is limited to gathering material and analyzing the probability scale, but also a philosopher-orator actively practicing the art of rhetoric. Describing one's results consists in translating the weighing of arguments into a text that by its nature has rhetorical power and itself is the result of applying mechanisms that rhetoric describes (Ricoeur, 2004, pp. 314–332).

Recounting the results of research, one produces a text of specific rhetorical power. This makes the way a given phenomenon is presented, and the positions one assigns the actors of a dispute, of key importance. In every presentation, someone begins a dispute and someone ends it; someone defends their cause and someone attacks it. The researcher must ultimately judge the arguments, present them as legitimate/unjustified, conclusive/inconclusive, accurate/inaccurate, etc.

Even if the values assigned in this description are formalized, i.e. they seem "technical" and related only to the structure of the dispute, they nevertheless involve making value judgments. Trying to limit this to logical coherence would be absolutely justified, but a coherent or incoherent argument, an accurate or inaccurate one, are arguments in favor of a given cause, and judging the argument indirectly affects judgment of the cause behind it.

If we are dealing with an abstract point of contention, it is easy to give readers a safe sense of impartiality. However, when we analyze disputes connected with religious thinking and religious worldviews, it becomes difficult to say the least. Of course this has its justification. Here is how Dilthey presented the problem:

> Nothing is more fleeting, frail, fickle than the mood of man toward the entirety of things where he finds himself, than the ideas he gets about the entirety of life and the world. Life either casts a dark shadow over our soul or lends it light and joy. How our days pass, how we feel life and where our experiences and notions fail us; the fact that we accept the world one day and reject it another, all this is affected by the way we perceive life. Nothing is more fragile than these images within us. Like light and shadow pass over a landscape, so does our sense of life change. Whoever wants to speak of human worldviews must be filled with it;
>
> Nichts ist flüchtiger, zarter, veränderlicher, als die Stimmung des Menschen gegenüber dem Zusammenhang der Dinge, in dem er sich findet, die Vorstellungen, die in ihm über den Zusammenhang des Lebens und der Welt entstehen. Das Leben wirft bald tiefe Schatten über unsere Seele, bald teilt es ihr Licht und Freude mit. Wie die Tage selber wechseln, wie wir das Leben fühlen und wo die Erfahrungen und Begriffe uns verlassen, wirkt unser Lebensgefühl darauf, wie wir den Zusammenhang der Dinge den einen Tag plausibel finden, den anderen verwerfen. Nichts ist zerbreitlicher als diese Bilder in uns. Wie Lichter Schatten über eine Landschaft geben, wechselt unser Gefühl des Lebens. Hiermit muß sich jeder erfüllen, der über die menschlichen Weltanschaungen sprechen will;
>
> (English translation based on original Dilthey, 1931, p. 168).

The above passage is rather poetic in character, probably because it does not come from the main part of the treatise *The Types of Worldview and Their Development in the Metaphysical Systems* (1911) but from the fourth addendum to it about religious worldview (*Zu: Religiöse Weltanschauung*). Nevertheless, it captures a fundamental feature of worldview structures and their perception by others. Worldviews are complex systems. Not only do they define fundamental rules of conduct based on a systemic image of the world (Dilthey, 1931, pp. 82–84), but their sources as well as the grounds for them stem from the attitude toward life of those who hold worldviews.

Our attitude toward life, meanwhile, cannot be reduced to the logical coherence of argumentation; it is the product of experiences, emotions, reasonings, encounters and probably many other, often elusive factors. It is impossible to conduct their uncontroversial stratification, nor is it possible to assess the percentage share of their influence on someone's worldview as a whole. Meanwhile, since disputes over religion result from the worldviews adopted by their participants, the disputes themselves cannot be presented in isolation from them. Describing a religious dispute, you describe not only the conclusiveness and coherence of argumentation, but also the ideas backing up individual arguments. This is where the worldview of the researcher starts to come strongly into play. How?

The hermeneutic trap

The issue here is not a simple expression of "I agree" or "I disagree," i.e. taking an unequivocal stand toward specific problems. You can stay aloof as a scholar in your studies, but I posit that you cannot do so a hundred percent consistently. At least that is what my own research experience has shown; its outline is the next point in my argumentation. I realize that the empirical nature of this argument demands caution during any attempt at generalization. Nevertheless, I would like to treat my own theoretical experience as a model of possible relationships between theoretical assumptions and the material being studied.

Setting about analyzing the dispute over Catholic modernism, I decided to conduct these studies from a hermeneutical perspective. Dilthey's observations on the task of philosophy and the nature of the subject of its study, namely worldviews, were

extremely important to me. However, I considered Gadamer's thoughts on the scholar's position with respect to the issue being studied. In the essay *Text and Interpretation* he wrote:

> Likewise, in the experience of history we find that the ideal of the objectivity that historical research offers us is only one side of the issue – in fact a secondary side, because the special feature of historical experience is that we stand in the midst of an event without knowing what is happening to us until in looking backwards we grasp what has happened. Accordingly, in every new present, history must be written anew. Ultimately, the same point holds true for philosophy and its history (Gadamer, 2007, p. 160).

Gadamer states outright that objectivity is of secondary importance in historical research because only a relative sense should be assigned to the notion. The researcher is always inside a historical experience, not outside it. No temporal distance changes that, meaning that describing segments of history is always an action performed in relation to a specific present. This enables thinking to always be in flux and causes nothing in the humanities ever to be proven once and for all.

Awareness of the temporal relativity of one's own results, and also of the direct connection between a researcher's interpretation and the context of the time and culture he or she lives in, is proof of self-awareness in the hermeneutical method as it realizes its own limitations. In a paradox, however, this kind of relativity clause can augment readers' sense of safety of reading and their level of objectivity. The fact is, unconcealed engagement is no longer a rhetorical surprise, but gives readers the possibility to distance themselves from it. If readers know where the author stands, they can separate and distinguish that standpoint from their own.

However, performing a post-factum self-analysis of the completed research process, I noticed that choosing the hermeneutical method had an enormous influence on the final shape of my research results, and remained in a strong relationship with my own worldview. Because, the point here is not the rather uncontroversial finding that the tools of interpretation this method offers impose a certain way of organizing the material. Rather, invoking hermeneutics enabled me to put forward the issue itself and, by presenting the dispute surrounding it, to boost those aspects that were part of the modernist position. Why?

To answer this question I need to start from the definition of Catholic modernism that – taking into account the tradition of research on the issue – I have adopted (Rogalski, 2018, pp. 21–26). It is a spectrum of worldview standpoints that share a belief in the necessity to reconcile Catholicism from the turn of the 19th and 20th centuries with contemporaneous civilization. Despite its antimodernist opponents, this movement cannot be reduced to one cohesive philosophical system. On the contrary, the modernists formulated very different worldviews; however, all of them contained this important directive for conduct: reconciling Catholicism with the present.

Presenting the dispute between modernism and antimodernism on issues of key importance for Christianity such as conscience, the significance of the Bible and dogmas, I tried not to openly favor any side. I reconstructed the arguments behind the mistrust of contemporary times as well as those that backed the opposite attitude. However, my sympathies – expressed in the structure and the choice of the issue itself – were with the modernists. What else, if not sympathies, can you call recognizing this particular issue from the history of Christian thought as being important and underappreciated, and the very fact of studying it – useful? What else is a search for the continuation of modernism and its contemporary consequences?

This attitude was not unrelated to the chosen hermeneutic method. Since hermeneutics puts the categories of understanding and interpretation at the focus of its consideration, since it sees value in the need for constant updating of historical reflection in relation to the present of successive hermeneuts, then does this perhaps affect one's attitude toward modernist thought which assumes a necessity for similar openness and multiplicity of interpretation? I consider this relationship to be distinct. If I as a researcher set great value by the possibility of analyzing historical issues in the name of deepening the contemporary debate or in the name of upholding values such as open-mindedness and multiplicity, which a hermeneutical sensitivity implies, then this has to influence my attitude toward a historical stance that holds similar values in high regard.

I do not think there is anything wrong or "unscientific" about this kind of relationship. On the contrary, it does not undermine the ideal of disinterestedness in research. It only makes us understand it differently. Disinterestedness does not mean complete

separation from the current cultural, social and economic context in studies on any issue. The possibility of such complete separation is something I question here. Disinterestedness, rather, means consistently following a certain conglomeration of values stemming from one's adopted worldview, and conducting research in accordance with the rules one accepts, and not shaping the results in the interest of any direct economic or social advantage.

Dialectics of distance

The relationship between worldview, method and results also makes us look a little differently at the category of the researcher's distance. I do not think this concept is starting to no longer make sense, but its importance is definitely being relativized. Consistent and unchanging distance in research is impossible to maintain, because the very choice of one's subject of study testifies to a specific attitude toward the world and results from a certain worldview. The connection between a researcher's worldview and that person's research may not be easy to notice at first, but can definitely be described after the research reaches a certain stage, by means of an analysis of the rhetorical structure of the results.

Being aware that such a connection exists, however, is key for continual work on maintaining a researcher's distance. Distance in this case assumes a dialectical character. It has to be acquired every time at the stage of formulating ideas and making them materialize in the form of a cohesive text. Awareness of the rhetorical nature of one's own text opens the way to its analysis already when it is being produced, and enables it to be made disinterested. This just means that the researcher deliberately decides what is important to him or her, and directly indicates such commitment. This kind of research stands a chance of contributing not only to understanding the subject undergoing study, but also to understanding oneself.

Perceiving this duality makes us also look a little differently at the importance of the hermeneutical tradition. Analyzing Dilthey's oeuvre and the hermeneutical tradition originating from it, Elżbieta Paczkowska-Łagowska notes that the opportunity created by hermeneutics consists in accepting the primacy of historical reasoning, understood as reflection on human history that is essential to understanding humans (Paczkowska-Łagowska, 2012, pp. 38–49). In this approach, history is of fundamental

anthropological importance. It is reflection on history that gives us knowledge on who humans are and how they change.

If, however, we become aware of the link between hermeneutics itself and a specific system of values that one adheres to when choosing it as one's method of interpretation, then that knowledge which comes from hermeneutical reflection is circular in nature. History cannot be separated from the historical acting subject. Nothing is fully clear here. Someone looking in a mirror does not simply want to see someone else; instead and above all, they try to see themselves.

References

Aristotle (1991). *On Rhetoric: A Theory of Civic Discourse*, transl. G. Kennedy. Oxford: Oxford University Press.

Arnold, C. (2007). *Kleine Geschichte des Modernismus*. Freiburg im Breisgau: Herder.

Dilthey, W. (1931). *Wilhelm Diltheys: Sesammelte Schriften*. Vol. VIII. Leipzig: Teubner.

Gadamer, H.-G. (2007). *The Gadamer Reader: A Bouquet of the Later Writings*, transl. R. Palmer. Evanston, IL: Northwestern University Press.

Kołakowski, L. (1988). *Metaphysical Horror*. Oxford: Basil Blackwell.

Koselleck, R. (1985). *Futures Past: on the Semantics of Historical Time*. Cambridge, MA: MIT Press.

Paczkowska-Łagowska, E. (2012). *O historyczności człowieka* [The Historicity of Man]. Gdańsk: słowo/obraz terytoria.

Ricœur, P. (2004). *Memory, History, Forgetting*, transl. Kathleen Blamey & David Pellauer. Chicago: University of Chicago Press.

Rogalski, M. (2018). *Producenci margaryny? Marian Zdziechowski i polski modernizm katolicki* [Producers of Margarine? Marian Zdziechowski and the Polish Catholic Modernism]. Kraków: TAiWPN Universitas.

White, H. (1973). *Metahistory: The Historical Imagination in Nineteenth--Century Europe*. Baltimore: Johns Hopkins University Press.

White, H. (1987). *The Content of the Form: Narrative Discourse and Historical Representation*. Baltimore: Johns Hopkins University Press.

Rafał Zawisza
ORCID 0000-0002-8771-1408
Faculty of "Artes Liberales", University of Warsaw
»Artes Liberales« Academy

Hannah Arendt's Marranic Evasions and the Truth of Her Cryptotheology[1]

Abstract

The text is a re-reading of *The Human Condition* by Hannah Arendt with emphasis on the concept of natality, showing connections between this book from 1958 and the author's dissertation written three decades earlier. The linking point is the thought of Augustine of Hippo, whose topics Arendt re-wrote into modern language, aiming to ultimately overcome gnosis, of both Christian and modern provenance. The era of nihilism as a modern type of gnosis reduced life to a biological process, a fact Arendt counterbalanced with "the divinity of birth" and, more precisely, with the latent sacredness of spontaneity. However, the operation of overcoming gnosis can only be grasped through careful research on Arendtian language, which is full of nuances. The hypothesis that Arendt worked on concepts as a "philosophical Marrano" helps elucidate the assertion that her secular anthropology is at the same time a cryptotheology.

Keywords: Hannah Arendt, Augustine of Hippo, philosophical anthropology, natality, birth, gnosis, secularity

Whereas in *The Origins of Totalitarianism* (1951) Hannah Arendt offered an analysis of recent history, in *The Human Condition* (1958) she undertook a much longer historical journey

to explain the crisis in contemporary politics and how we can overcome it. Thus, the two books can be read as a diagnosis and as a proposal, respectively. *The Human Condition* did not appear out of a void. Arendt started to think about natality again just as soon as she finished the manuscript in 1949. When she understood, for herself, what had happened, her decision was to construct a *political anthropology* which could provide an answer to the challenges of post-totalitarian politics. One dimension of her work consisted in *smuggling old meaning into new concepts*, which did not leave this meaning intact. This was done against tradition, which she regarded as broken and finished. Hannah Arendt's cryptotheological intervention was intended to *regain a certain conception of life not inherited via tradition*. Hence, ossified political and theological concepts were avoided. Since theology was finished too, Arendt's intellectual operation was not to rescue politics by grounding it in some theology; it was rather an attempt to excavate the most ancient and most forgotten ideas in order to give them a new life.

* * *

The Human Condition opens with the incredible *Prologue* in which Arendt convincingly depicted what she felt determined the spiritual atmosphere of the new age. A massive unspoken *Grundbefindlichkeit*, which had been formed during the two most devastating wars history had ever known, found its expression when the first "earth-born object" (HC: 1) was sent into the cosmos. Arendt noticed a dominant feeling of "relief" that humankind would no longer be imprisoned on Earth (HC: 1). That would mean that the planet presented itself as an unsafe place from which one would like to escape. There is no doubt that a great deal of the author's own feeling was manifested in this image, since those who had to flee their own countries, and were unable to find any other safe place to live, could have perceived the world as radically alien and hostile. Claustrophobia of this sort was something unprecedented for Arendt. That may well be so when it comes to its scale, but there was at least one other well-known historical formation which shared similar basic impressions with the inhabitants of the postwar West: the ancient Gnostics. And, there is a certain Gnostic imaginary accompanying *The Human Condition*. Whereas Eric Voegelin characterized the modern age as the Gnostic era, Hannah Arendt

took up this idea, but in her writing, it swerved. Accusation gave way to existential description:

> Should the emancipation and secularization of the modern age, which began with a turning-away, not necessarily from God, but from a god, who was the Father of men in heaven, end with an even more fateful repudiation of an Earth who was the Mother of all living creatures under the sky? (HC: 2)

The dualist system of two deities, as found in some variants of Gnostic theology, protects the image of good, true God from being conflated together with the bad reputation of "a god" who is not the true God and who has to wane precisely for that reason. By distinguishing between God and a god, Hannah Arendt alluded to the living God of Israel, who – unlike its philosophical simulacrum – can live on despite secularization. Interestingly enough, God the Father here appears as a mistaken image, even though theology usually has no objection to refer to it. It is certain that this protective and sovereign deity was known to pagan cults as well as to the Jewish and Christian religions. Somehow, however, Arendt did not treat this vision of a potent, powerful divinity as essential; perhaps it was more reasonable and understandable when humanity lived in misery and needed a mighty patron.

Either way, we do not get to know much about the true, living God: The accent is put on the other pole of religious piety focused on the Earth and condensed in the image of the Mother. Identified with what is "under the sky," what our earthly reality is built of, the metaphor of the Mother could be more substantial and tangible. If there is something dangerously capricious in the figure of the sovereign Father, the Mother-Earth provides unquestioned affirmation:

> The earth is the very quintessence of the human condition, and earthly nature, for all we know, may be unique in the universe in providing human beings with a habitat in which they can move and breathe without effort and without artifice. The human artifice of the world separates human existence from all mere animal environment, but life itself is outside this artificial world, and through life man remains related to all other living organisms (HC: 2).

However sweet the maternal influence on human life might be, it cannot enclose it in a fully protective bubble. Partially

safe in the bubble of civilization, human beings are exposed to participation in the stream of Life, which also includes a violent fight for survival, as well as death. Hence, the image of Mother will always remain ambivalent[2]. In addition, what makes the Gnostic language more prominent is that human life not only juts out of the oikumene, but it also juts out of the Earth overlooking the cold cosmos[3]. We live both inside and outside of the bubble.

"The wish to escape the human condition" springs from a fear of inhospitality, and prompts man to create life artificially, "cutting the last tie through which even man belongs among the children of nature" (HC: 2). The phrase "even man" signifies that this is a bizarre fact, not at all self-evident and perhaps even not easy to accept. Although Arendt did not claim that this represented a return to some kind of natural law, she nonetheless insisted on respect for the basic conditions that determine not only human life, but life on the planet in its entirety. Arendt was skeptical about the future possibilities of humans to shape their own destiny and their own organic life, but she admitted that "science has realized and affirmed what men anticipated in dreams that were neither wild nor idle" (HC: 2). In different circumstances, she surmised that the progress made by science would have been celebrated with "joy" and "pride," even a feeling of triumph (HC: 1), but what scared her was "a rebellion against human existence as it has been given, a free gift from nowhere (secularly speaking)" (HC: 2–3). Again, this is not to say that humans should stick to a pre-determined shape of existence, but they should respect the limits within which this condition is "given."[4] It is interesting to know that even "secularly speaking," one can still speak about a "gift." Gift, according to Arendt, does not imply

---

[2] I would risk the hypothesis that Kristeva's intellectual portrait of Arendt (Kristeva 2001) is so excellent because it is preceded by a deep philosophizing on the works of Sigmund Freud, Melanie Klein, and Jacques Lacan, whose vision of the mother figure was extremely ambivalent.

[3] Arendt compared oikumene to a house, using apolitical language: "Work provides an 'artificial' world of things, distinctly different from all natural surroundings. Within its borders each individual life is housed" (HC: 7). It was precisely this pre-political and apolitical language that she identified in the *Concluding Remarks* from *The Origins* as the founding idiom of human rights.

[4] In the text, Arendt repeated her theoretical early gesture: Cutting the umbilical cord that used to connect us to God implies that there is no reason for nostalgia.

the existence of a giver. It is a *metaphor* which operates in the register of *secular sensitivity with a Gnostic touch*: The attitude towards the gift of life could turn into a negationist resentment that was alien to her. Arendt emphasized that we are "earth-bound creatures and have begun to act as though we were dwellers of the universe" (HC: 3). Not necessarily divine creatures, or: even if divine, we are first of all bound to the Earth and responsible for this reinvigorating bond. The response which seemed adequate in Arendt's eyes was *gratitude*, which reflects the affirmative appreciation of the life given.

The category of creatureliness served Arendt not only to reinforce connectedness to the world – contrary to the vector of the crushing majority of traditional theological production – but also to protect the equality of men. Once again, Arendt used the initial pages of *The Human Condition* to engage theological sources in order to combat theology and to propose her own reading of Genesis. To create her own philosophical anthropology, she decided to discuss the origins of humankind and the importance of plurality, but first she needed to contest very old habits of reading the Bible. It might be surprising to a secular reader that after a brief note about the very political understanding of life by the ancient Romans, Hannah Arendt went back to the biblical motifs. More space is devoted to the Judeochristian tradition[5], even though in the whole book the Greek (and Roman) heritage is at the center of attention. Yet, it was not the Greek or Roman founding myths, but the biblical story about the Creation that Arendt found more capable of transmitting a truly universalist message.

Arendt claimed that Jesus of Nazareth was *the* true teacher of action, and she saw the whole Judeochristian tradition as supporting an augmented historical agency, since "in its most elementary form, the human condition of action is implicit

---

[5] I use this word without a hyphen deliberately, since Hannah Arendt belonged to a group of Jewish thinkers whose intellectial activity operated precisely in the cultural space where that hyphen makes no sense. It is not to negate the differences between Judaism and Christianity, but to respect the experiences of those extraordinary thinkers who thanks to their religious commitment or secularity did not respect the divisions created during centuries of living together and apart. For an interesting study on the emerging intra-confessional boundary, see the work of Daniel Boyarin (Boyarin 2004). See also a chapter by Itzhak Benyamini (Benyamini 2016).

even in Genesis ('Male and female created He them')" (HC: 8). It would be wrong, however, to state that this tradition always stood for freedom and plurality. Above all, what was always somehow excluded from the theological landscape, apparently seen as scandalous, is the simple fact that "the multitude of human beings becomes the result of multiplication" (HC: 8). It was not always like that, as the controversy surrounding human sexuality was emphasized significantly after Paul became the leader of Christianity. It was he – Arendt noted in the first footnote of the book – who preferred the version of the Creation from Genesis that insisted on the primacy of man over woman. Jesus, instead, accentuated the inborn plurality of men: "The difference indicates much more than a different attitude to the role of woman. For Jesus, faith was closely related to action …; for Paul, faith was primarily related to salvation" (HC: 8)[6]. Not only that: Later on, it was a very political decision every time someone had to choose between two competing biblical versions of the Creation. Thus, Arendt admitted the importance of religious heritage for "postclassical political thought" (HC: 8), and from the beginning showed how even within the original text – not to mention apocryphal literature or heterodox theologies – there were seeds of different interpretations. It is not so much the holy text itself as rather the decisions of its interpreters which should be judged. Having dealt with gender issues, Arendt turned her gaze to the most precious thing which Judaism created – singularity:

> Especially interesting in this respect is Augustine (*De civitate Dei* xii. 21), who not only ignores Genesis 1:27 [the version emphasizing equality – RZ] altogether but sees the difference between man and animal in that man was created *unum ac singulum*, whereas all animals were ordered "to come into being several at once" (*plura simul iussit exsistere*). To Augustine, the creation story offers a welcome opportunity to stress the species character of animal life as distinguished from the singularity of human existence (HC: 8).

---

[6] Since in their "Dual Monarchy" Heinrich and Hannah thought together, it is worth noting what Blücher stated in one of his lectures at Bard. According to him, Jesus "established the equality of human beings in quality, in the infinite possibilities of every person, and therefore in the absolute inviolability of that person. But he knew that in order to establish this he had first to abolish the inequality between man and woman, because here every other form of inequality was anchored" (Young-Bruehl 2004: 269).

Singularity is the birth canal of plurality. Augustine conceptualized them as opposing poles, Arendt as being interrelated[7]. The majority of interpreters ignore the importance of singularity completely; in that, they are a symmetrical semblable of Augustine – they just change "plus" to "minus." Hannah Arendt was more theologically educated and took care to be cautious in her critique. She regarded the revulsion felt towards plurality as unacceptable from an anthropological as well as a theological point of view. On the contrary, human isolation, if not seen as absolute and autarkic, is a matter of fact because of mortality and temporality of thinking. Augustine accommodated a poisonous mixture of Manichaean and neo-Platonic principles in the Judeochristian tradition, which made his philosophy so tragically dramatic, contradictory and thought-provoking. However, this was the very same reason that led him to formulate the doctrine of original sin, which ultimately was supposed to pacify tensions and eliminate elements dangerous to ecclesiastical power. The price for this neat reconciliation – which, however, provoked many more dilemmas than it managed to resolve[8] – was a negative view of the human body, its desires, and reproduction.

No wonder the Augustinian legacy is discussed on the same page in *The Human Condition*, and for the first time the word "natality" appears. Natalism, Hannah Arendt's highly novel version of vitalism, is her answer to the conception of original sin. It is a cryptotheological critique which builds its own consistency in a polemical mode, which allows for appropriation and creative misreading. Arendt presented her anthropology by commenting on Augustine because he was the culmination of the two most powerful traditions behind Western civilization: Judeochristian and Graeco-Roman. Moreover, through the polemics which Augustine had with his contemporaries, he unintentionally preserved remnants of less influential cultures and religions, especially the

---

[7] Arendt speaks in parenthesis of the "perfect 'singularity'" of a philosopher contemplating eternity, because this view of singularity "can occur only outside the realm of human affairs and outside the plurality of men" (HC: 20). This view is misguided by the false ideal of autarky. In that form, Arendt's critique does not turn against singularity, but against autarky.

[8] That is why Hans Blumenberg stated that the Augustinian answer to Gnosticism and his attempt to overcome the latter was unsuccessful (Blumenberg 1985). It was modernity that took on the task of a second attempt at overcoming Gnosticism.

legacy of various Gnostic groups as well as other theological alternatives, which orthodoxy resisted.

The use of the polemical mode is important because cutting ties with somebody always means that the emancipated party will carry along the principle of its independence. Independence could be real and effective, yet the way it comes about gives the new phenomenon a specific flavor. This is why I call Arendt's anthropology cryptotheological – not only because it was born during her study of theological tractates, but also because it carries an *intentionally distorted message*. The secularity of Arendt's proposal has a visibly theological, not to say divine, origin. For Christians (and for thinkers sticking to the Christian legacy, like Schmitt and Voegelin) this means that secularity is dependent on its religious past and every attempt to emancipate from it is nothing but rebellion. For Arendt, this is not a foregone conclusion: Jews never had such a problem with worldliness as Christians did. For Jews, the Creation was never evil, whereas Christians risked this possibility precisely in the conception of original sin, meant to relegate the Manichaean threat, but which instead incorporated a negative perception of matter (and reproduction) while still being obliged to maintain the goodness of the Creation as the basic premise. Christian theology needed a scapegoat and there could be no better candidate than humans.

This reconstruction of the psychotheological motives behind man's condemnation, presented by Hans Blumenberg, finds an exact counterpart in Arendt's writings. It would be seemingly contradictory that in *Augustin und der Protestantismus* (1930) she claimed that singularity was of no interest to Augustine, while in *The Human Condition* (1958) singularity, as we have seen, is a hallmark of man. Did she change her mind? Not really. In 1930, she noted that singularity was not essential to Augustine *as such*, as an emphatic sign that every human being is unique.

In the Church Father's perspective, man should be seen first of all in the light of salvation, not in his earthly, transient form. A weakness for mundane charms is typical of *cupiditas*, the sinful lifestyle. So, if singularity was not attractive to Augustine the Christian, why did he make it central in his conception? The answer is because it perfectly suited the pressing demand that arose from the question of who is responsible for evil. The question was posed on a cosmological scale, and thus the guilt imposed on humans was given cosmic weight. Augustine's "discovery" of

deep inwardness and singularity was simultaneous with its "creation" – in other words, burdening humans with cosmic responsibility paradoxically carved out more space for their exceptionality. The consciousness of the individual became the necessary place where the entire drama could be staged. It was not until the modern rediscovery of human personality that the singular reappeared. But one has to concede that it was not Augustine alone who invented singularity *ex nihilo*; he made it irreversibly central and laden with tragic meaning, but it had been essential to Judaism centuries before, always waiting for re-activation.

Arendt's revival of *singularity within plurality* – in accordance with the *secular type of transcendence within immanence* – seems to be at once Judaic and modern, if compared with the theological reservations of Christianity towards reproduction. In addition, there is a strong antinaturalistic and anti-cosmic attitude in Arendt:

> Action would be an unnecessary luxury, a capricious interference with general laws of behavior, if men were endlessly reproductible repetitions of the same model, whose nature or essence was the same for all and as predictable as the nature or essence of any other thing. Plurality is the condition of human action because we are all the same, that is, human, in such a way that nobody is ever the same as anyone else who ever lived, lives or will live (HC: 8).

What has always struck me in the ideal vision of life presented by traditionalists is that freedom is held not only as a pernicious interference, but also as a luxurious caprice of nature. And indeed, human beings are – by nature – unnatural; we are *freaks of nature*. If one does not accept this simple fact, then things become very problematic, because culture or politics which is anti-factual stubbornly resists the human essence in its expression and cultivation. To blame humans for being excessive is to commit a cardinal mistake, because we are born this way and only strong cultural repression could remove this fact from sight. Nevertheless, meditation on birth allows a reversal of the eclipse of the amazing energetic potentiality symbolically and factually condensed in birth. Most of all, a new anthropology needs to unlearn certain mental habits resulting from the resonation of the doctrine of original sin. Humans are not evil, not even dangerous at the core – humans are first of all afraid of their own existence and mortality.

At first, Arendt introduced the most general determinants of the human condition as equal: "birth and death, natality and mortality" (HC: 8). They seem to be symmetrical, yet we have to remember that Arendt was not a traditionalist for whom harmony and symmetry, or some other aesthetical or mathematical measures, could dictate what matters for human life. For this reason, the significance of natality immediately comes to the fore:

> Labor and work, as well as action, are also rooted in natality in so far as they have the task to provide and preserve the world for, to foresee and reckon with, the constant influx of newcomers who are born into the world as strangers. However, of the three, action has the closest connection with the human condition of natality, the new beginning inherent in birth can make itself felt in the world only because the newcomer possesses the capacity of beginning something anew, that is, of acting. In this sense of initiative, an element of action, and therefore of natality, is inherent in all human activities. Moreover, since action is the political activity par excellence, natality, not mortality, may be the central category of political, as distinguished from metaphysical, thought (HC: 9).

It is often repeated in scholarship what Arendt wrote about action: It is ontologically grounded in the fact of birth, thanks to natality. But it is almost never mentioned what she stated earlier, in a less philosophical idiom – namely that labor, work, and action are "intimately connected" (HC: 8) with birth and death. This could mean that by working, creating things and acting, humans take into account their finitude, but above all they have a creative and disruptive energy. Mortality is in the shadow when there is a place to realize natality. On the other hand, when the capacity of initiating action is suppressed, then life falls under the mark of death, through adaptation to social demand, which for a free being is akin to death. This is a killing of the potential for renewal. In the above passage we see the great change in Arendt's thinking that occurred between 1951 and 1958. In *Concluding Remarks*, she started to think about birth politically, but it was not yet a complete philosophical anthropology. Moreover, there she claimed that the commemoration of the victims of genocides will stand at the heart of any future political thinking. *The Human Condition* thus brought about a decisive moment: The promise of birth that shone through *Ideology and Terror* finally gained its advanced articulation.

The cryptotheology of singularity which preceded Arendtian political theory was itself preceded by the anthropogenetic narrative built on recent discoveries of the natural sciences. This peculiar theoretical construction – the concept of natality floating between theology and biology – mirrored the layers of thinking which one may find in the author Arendt referred to, who had reflected upon the individual, worldly, and cosmic perspective – namely Augustine. No wonder her description of the human condition and its earth-bound character derives from the theology of God's enigma.

> The problem of human nature, the Augustinian *quaestio mihi factus sum* ("a question have I become for myself"), seems unanswerable in both its individual psychological sense and its general philosophical sense. … Moreover, nothing entitles us to assume that man has a nature or essence in the same sense as other things. In other words, if we have a nature or essence, then surely only a god could know and define it, and the first prerequisite would be that he be able to speak about a "who" as though it were a "what." The perplexity is that the modes of human cognition applicable to things with "natural" qualities, including ourselves to the limited extent that we are specimens of the most highly developed species of organic life, fail us when we raise the question: And *who* are we? This is why attempts to define human nature almost invariably end with some construction of a deity, that is, with the god of the philosophers, who, since Plato, has revealed himself upon closer inspection to be a kind of Platonic idea of man. Of course, to demask such philosophic concepts of the divine as conceptualizations of human capabilities and qualities is not a demonstration of, not even an argument for, the non-existence of God; but the fact that attempts to define the nature of man lead so easily into an idea which definitely strikes us as "superhuman" and therefore is identified with the divine may cast suspicion upon the very concept of "human nature" (HC: 10–11).

This is a classic example of Rosenzweig's new thinking (Rosenzweig 2005), where anthropological description creates a collision of philosophy and theology – here not even as competing discourses, but as totalizing ones which equally eclipsed the mystery of the intimate relation between man and God. First and foremost, the concept of human nature served to instrumentalize man, to find some regularities which instead of freedom, concentrated on what is necessary. The reification or animalization of man served to overlook human dignity, which is impossible without that which crosses the boundaries of necessity.

There is a huge contrast between "the god of the philosophers" and the "God" discreetly lurking in Arendt's text. The intuition of a true God counterbalances all philosophical and theological thematizations of the concept of God. The true God, even if he knows what the true nature of man is, never reveals this knowledge to man. The inaccessibility of human nature protects the human enigma. What is more, this enigma is intimately connected to the secret residing in God. Only God could know who we really are. "What" we are is known to the natural sciences, partly to psychology and sociology. But the very important shift which makes *Arendt's cryptotheology indistinguishable from secular thinking* is that she did not allow the enigma to reside solely in some depths of inwardness. "Who" we are, although it can never be exhausted and determined, is communicable between people. The whole construction of the public sphere, of speaking and acting, is based precisely on her cryptotheology. By expressing ourselves, we share something of this enigmatic dimension, mostly visible through the body's charm, idiosyncrasy of voice, gesture, and facial expression. *Expressivity is an allusion to the enigma, not its expression*[9].

In a footnote – the second in the book – Arendt referred to Augustine in detail, confirming not only her commitment to his legacy, but also stating quite boldly that "he knew this quite well" (HC: 10), whereas science still does not today, that human nature, even if it could be claimed to exist, is not detectable to scientific radars. Man is, to Augustine and also to Arendt, *grande profundum*, "the great mystery" (HC: 10). Since we are dealing with cryptotheology, its deepest convictions should have remained unsaid or else been encrypted in quotations from other authors. Arendt cited Augustine precisely in the vein of mystery so that the way the enigma of man exists would be exposed. A quotation from the *Confessions* is preceded by two words of Arendt's which intertwine with it smoothly, supporting my reading, namely that Arendt's voice is hidden here behind Augustine's: "there is 'something of man [*aliquid hominis*] which the spirit of man which is in him itself knoweth not. But Thou, Lord, who has

---

[9] In a letter to Mary McCarthy dated August 8, 1969, Arendt wrote that it is not expression but speech which distinguishes us from animals. What is equally important, Arendt did not think that speech, gestures, and facial expressions are a pure expression of inwardness (AM: 242–243).

made him [*fecisti eum*] knowest everything of him [*eius omnia*]'
(x. 5)" (HC: 10–11).

Apart from the question of how this vision of God shapes
anthropology, it is evident that the influence is mutual: "The
question about the nature of man is no less a theological question
than the question about the nature of God; both can be settled
only within the framework of a divinely revealed answer" (HC:
11). God is involved with anthropology, but – reversely – it also
affects the way divinity is perceived. In this way, one can legiti-
mately say that because of the chiasmus between God and man
in Arendt's thought, her political anthropology is at the same
time a cryptotheology.

* * *

To conclude, I will demonstrate an essential sample of Arendt's
Marranic style on full display, i.e. through an example of Judeo-
christian material. Ambiguity seems to be crucial in speaking
about the phenomenon of philosophical Marranism (Bielik-Robson
2014). If it is unclear whether one is talking about the Chris-
tian or the Jewish Messiah, the question of messianism acquires
a new dynamic. However, blurring divisions is as important as
the necessity of differentiation. Marranism does not constitute
syncretism, it is a careful deconstruction.

From there, a certain promise kept following Arendt's writ-
ings. She was aware that to preserve hope for a revolutionary
passage to a renewed humanity, it is necessary to abstain from
identifying this hope with any actual event or person. Already
before the war she warned her husband, Heinrich Blücher,
against identifying religious symbols with political reality. In
a long letter, he wrote to her that the Jews needed to abandon
the bourgeoisie and join the world revolution. He said Marx was
a prophet, he – Heinrich – called himself the "Wunder-Rabbi,"
and the figure of a Bolshevik would become the Messiah. Arendt
responded with unhesitating rejection of any such idea (WFW:
14–19). Nevertheless, she did preserve some messianic undertones
in her writings, which can be clarified through her Marranic
lecture on the Nativity.

Messianic hope works as a transcendent *reminder*. Such
a reminder was fully *incarnated* in Arendt's *text* and is con-
densed in a quotation without reference which she used in *The
Human Condition*, namely: "'A child has been born unto us'"

(HC: 247), which in the German edition was written as "'Uns ist ein Kind geboren'" (VA: 243). Commentators have argued about its origin and its meaning. It appears as the last sentence in the crucial chapter about action. Arendt ends this chapter by describing human finitude as living out of nature, depending on "Fähigkeiten des Neubeginnens" (VA: 242)[10], the faculty which is "like an ever-present reminder that men, though they must die, are not born in order to die but in order to begin" (HC: 246)[11].

Later Arendt spoke about Jesus as the teacher of action and new beginnings which are miracles "within the reach of man" (HC: 247). Some of the phrases are not to be found in the English version, like the clear statement that what interested the author was "die spezifisch politisch-philosophische Bedeutung der Geschichte Jesu" (VA: 243). The figure of "Jesus of Nazareth" – as he is called in *The Human Condition* – emphasizes what the Judeochristian tradition introduced and what was missing in Greek antiquity: faith and hope. The significance of Jesus in *The Human Condition* and *On Revolution* prompts us to scrutinize further the role of the ambiguous quotation. The final passage in the German version reads as follows: "Daß man in der Welt Vertrauen haben und daß man für die Welt hoffen darf, ist villeicht nirgends knapper und schöner ausgedrückt als in den Worten, mit denen die Weihnachtsoratorien 'die frohe Botschaft' verkünden: 'Uns ist ein Kind geboren'" (VA: 243).

In the English "original" we read about "the few words with which the Gospels announced their 'glad tidings'" (HC: 247). That would suggest Arendt took the word directly from the Bible. But we know from her correspondence that she collected these words during one of her trips to Germany, in 1952, where she listened to *Messiah* by Georg Friedrich Händel[12]. In the libretto,

---

[10] In the English version it is called "the faculty of … beginning something new" (HC: 246).

[11] The German version contains a crucial addition to this phrase, which allowed me to return once again to the concept of a human "substratum": the faculty of making beginnings is the one "die anzeigt, daß Menschen zwar sterben müssen, aber deshalb noch nicht geboren werden, um zu sterben, sodern im Gegenteil, um etwas Neues anzufangen, solange der Lebensproceß das eigentlich personal-menschliche Substrat, das mit ihnen in die Welt kam, nicht zerriben hat" (VA: 242).

[12] Reporting her discovery about the greatness of birth, she wrote to Heinrich ending with her characteristic sense of humor: "Das Hallelujah liegt mir noch im Ohr und den Gliedern. Mir wurde zum erstenmal klar, wie großartig

there are two quotations, one from the Book of Isaiah, the other from the Gospel of Luke, which could have resonated in Arendt's memory. If one looks at the King James Bible, Arendt's "A child has been born to us" differs from "For unto us a child is born" (Isaiah 9:6) as well as from Luke's: "For unto you is born this day in the city of David a Savior, which is Christ the Lord" (Luke 2:11). Luther's Bible does not contain an identical phrase for the passage: "denn euch ist heute der Heiland geboren, welcher ist Christus, der Herr, in der Stadt Davids." However, in Luther's translation of the Book of Isaiah we find: "uns ist ein Kind geboren," exactly as Arendt wrote in *Vita activa*. What does it all mean? Frederick M. Dolan says that "Arendt's sly rewriting of the Bible makes bad scholarship, but good sense" (Dolan 2004: 610). But is that not a bit banal? Was she not educated enough to know how to quote? Didn't she have a Bible at home to check the passage? I think, rather, that we are dealing here with what Franz Rosenzweig called the "mosaic style," which used to be a typical style of Jewish speech in which the level of absorption of religious texts was so advanced that the Bible lived in distorted crypto-quotations. Those who spoke in the "mosaic style" did not care about the accuracy of "citing"; in fact, to them it was not citing but their own parlance. Crypto-quotations constituted the very fabric of "the mosaic style."[13] I agree, though, that from a scholarly point of view, it may look sloppy. All in all, the word *marrano* literally means "sloppy" (next to "filthy," "pig," and "rude man").

Nevertheless, Dolan is right about the double move Arendt made: She told us about each and every birth, not only about the appearance of Jesus Christ. Yet, thanks to that inaccurate quotation, she opened her text to Christian readings, thus universalizing the Jewish promise from Isaiah and, additionally, depriving it of the dimension of a future earthly kingdom of God (Dolan 2004: 608). However, the divinity of Jesus was also annulled. Jeffrey Champlin very nicely grasped the message hidden in Arendt's oscillation and creative misreading:

---

das: Es ist uns ein Kind geboren, ist. Das Christentum war doch nicht so ohne" (AB: 270) – "Christianity was not nearly so dumb"…

[13] See Marie Luise Knott's book (Knott 2013: 158), in which she makes a reference to Klaus Reichert and his *Die unendliche Aufgabe: Zum Übersetzen*, Hanser, München–Wien 2003.

From a conceptual point of view, we can see that in the inaccurate citation the out of place child finds the mis-place that was prepared for it. In this sense, Arendt shows that a birth never arrives where it should. In her aesthetic conceptualization, she secularizes the story of Christ's birth by showing that there must always be room for the child who has been announced but not expected (Champlin 2013: 164).

There is a point, however, on which I disagree with Champlin, who thinks that secularization means living in a world that is fully profane, without any trace of the divine. This understanding of secularization and of being secular does not apply to Hannah Arendt as a secular Jew and a philosophical Marrano. In *On Revolution*, commenting on the fourth Eclogue of Virgil, she replaced "the arrival of a divine child and savior" – traditionally identified with Jesus – with "the divinity of birth as such" (OR: 211). It is the only place in her whole published corpus where her cryptotheology becomes an *open secret*. It is like a *crater in the text*, left without comment. A reminder: telling and irritating.

References

AB   Arendt, H., Blücher, H. (1996). *Hannah Arendt/Heinrich Blücher: Briefe 1936–1968*. Herausgegeben und mit einer Einführung von Lotte Köhler. München: Piper.

AM   Arendt, H., McCarthy, M. (1995). *Between Friends: The Correspondence of Hannah Arendt and Mary McCarthy, 1949–1975*. Edited and with an introduction by Carol Brightman. New York: Harcourt Brace.

HC   Arendt, H. (1958). *The Human Condition*. Chicago: University of Chicago Press.

OR   Arendt, H. (1990). *On Revolution*. London: Penguin Books.

VA   Arendt, H. (1960). *Vita activa: Oder vom tätigen Leben*. Stuttgart: W. Kohlhammer Verlag.

WFW Arendt, H., Blücher, H. (2000). *Within Four Walls: The Correspondence Between Hannah Arendt and Heinrich Blücher, 1936–1968*. Edited and with an introduction by Lotte Kohler. Translated from the German by Peter Constantine. New York: Harcourt.

Benyamini, I. (2016). The hyphenated Jew: Within and beyond the 'Judeo-Christian'. In: E. Nathan, A. Topolski (Eds.), *Is There a Judeo-Christian Tradition?: A European Perspective* (pp. 225–240). Berlin: De Gruyter.

Bielik-Robson, A. (2014). *Jewish Cryptotheologies of Late Modernity: Philosophical Marranos*. London: Routledge, Taylor & Francis Group.

Boyarin, D. (2004). *Border Lines: The Partition of Judaeo-Christianity*. Philadelphia: University of Pennsylvania Press.

Blumenberg, H. (1985). *The Legitimacy of the Modern Age*, transl. R. M. Wallace. Cambridge, MA: MIT Press.

Champlin, J. (2013). Born again: Arendt's 'natality' as figure and concept. *The Germanic Review* 88(2): 150–164.

Dolan, F. M. (2004). An ambiguous citation in Hannah Arendt's *The Human Condition*. *The Journal of Politics* 66(2): 606–610.

Knott, M. L. (2013). *Unlearning with Hannah Arendt*, transl. D. Dollenmayer. New York: Other Press.

Kristeva, J. (2001). *Hannah Arendt*, transl. R. Guberman. New York: Columbia University Press.

Rosenzweig, F. (2005). *The Star of Redemption*, transl. B. E. Galli. Madison: University of Wisconsin Press.

Young-Bruehl, E. (2004). *Hannah Arendt: For Love of the World*. Second edition. New Haven: Yale University Press.

Piotr Kałowski
ORCID 0000-0001-9588-3923
Faculty of Psychology, University of Warsaw

Narration True and False: Dialogical Self Theory in Psychotherapy

Abstract

Postmodernism denies universal values and emphasizes the role of political power in constructing truth. This creates a challenge for psychotherapy, where the client's perspective must simultaneously be respected and filtered through a given modality's theory. The paper discusses the notions of truth and falsity in psychotherapy, presenting dialogical self theory as an approach responding to the aforementioned challenges. Comparing it to the dominant cognitive-behavioral approach, examples of its practical use, especially problems related to identity and its continuity, are described, and the benefits and limitations of the metaphor of therapeutic work as a dialogue or narration are outlined.

Keywords: Postmodernism, dialogical self theory, narration, psychotherapy

> I seek the truth about the other, meaning agreement
> between what is, and what is being said.
> Opoczyńska (2012), p. 371

Introduction

A popular joke among psychologists involves responding "It depends" to questions about the science. Indeed, many psychological phenomena are described by several coexisting theories, and studies often report contradictory results. Thus, emphasis is increasingly being put on the relativity of knowledge. The answers we receive depend on the questions we ask, and the act of observing must necessarily involve the observer in the observed.

One area where this relation seems especially important is psychotherapy. Here, a major problem is that while its effectiveness

has been proven (American Psychological Association, 2013), the knowledge of *how* therapy works is, for the time being, incomplete. In other words, practice informed by theory brings results, but the theories have not been tested in their entirety (David & Montgomery, 2011). How, then, should the notions of *truth* and *falsity* be applied to situations where we know how a certain cause-and-effect sequence occurs, but we don't know why?

How many truths?

Emotional suffering has been with us since the beginning of our species, together with our attempts to alleviate it (Orlinsky, 2017). The evolution of these attempts thus parallels the evolution of our values: what is the norm, the ideal, the problem. This history is, in turn, made up of countless individual histories or *narrations* (Angus & McLeod, 2004). On both the global and the individual level, those narrations seek to make sense of this suffering (see Ghaemi, 2007), though history is replete with examples of tragic divergences between them (see, e.g., Showalter, 1985).

Today, over a hundred years since the collapse of the hegemony of psychoanalysis (Schultz & Schultz, 2008), it is estimated that there exist around 1,000 different theories or schools of psychotherapy (Garfield, 2006). Thus, the research question is not "Is therapy effective?" nor "Which therapy is the most effective?" but rather "What treatment, by whom, is most effective for this individual with that specific problem, and under which set of circumstances, and how does it come about?" (Paul, 1969, p. 44).

One proposed answer is the *evidence-based psychotherapy* (EBP, see Hayes & Hofmann, 2018) paradigm. Carried over from medicine, it posits informing clinical decisions to the greatest possible extent with results of *randomized controlled trial* (RCT) studies, which mirror clinical pharmacotherapy trials. However, since RCTs require strict control of factors related to the person of the client and the therapist as well as the therapy setting, they often involve specific methodological limitations (Grant, 2009). Consequently, elevating them to the "gold standard" has met with criticism, as it reduces the complex therapeutic relationship to a set of bullet points (Koerner, 2018). As EBP narrows the definition of acceptable scientific evidence, it becomes "dangerously reductive insofar as it negates the personal and interpersonal significance

and meaning of a world that is first and foremost a relational world, and not a fixed set of objects" (Holmes, Murray, Perron, & Rail, 2006, p. 183). Overconfidence in RCT results, together with the limits of insight into one's own decision-making processes (see, e.g., Kozlowski, Hutchinson, Hurley, Rowley, & Sutherland, 2017), thus creates a risk of rigidly fitting clients' narrations into the "correct" therapeutic blueprint.

Towards multivoicedness

An interesting response to EBP's perceived shortcomings is the *narrative medicine* approach (Charon, 2012). Focusing on the patient–doctor relationship rather than just on the symptom–treatment correspondence, narrative medicine encourages "acting on the realization that a portion of the salient evidence will be found not in numbers but in language" (Goyal et al., 2008, p. 733). By being sensitive to power and perspective differences (e.g., the patient's mistrust of pharmacotherapy) and by informing the patient about the arguments behind their clinical decisions, practitioners can gain insight into the personal meaning the symptoms have for their patients. In turn, considering these symptoms through the lens of the patients' life narratives enables more effective care to be provided via a dialogue (Charon, 2011). The patient thus ceases to be a clinical case to be "cracked" according to the rules of the dominant theory. Rather, the clinician "admits all possible light instead of just the most comfortable or convenient light" (Charon, 2012, p. 7).

On the other hand, by fostering the understanding of "the clinical consultation as a story, set against a complex backdrop of personal history, culture, ethnicity, gender, and economic status" (O'Mahony, 2013, p. 614), narrative medicine creates the risk of dangerous emotional entanglement with the patient, which, under the guise of empathy, blurs responsibility for decisions about treatment.

Such an approach seems more adequate in psychotherapy, where the emotional level melds with the intellectual, and where clients' narrations are often the source of symptoms (Machado & Goncalves, 1999). Taking a constructionist stance, *narrative psychotherapy* states that the fundamental element of human mental activity is organizing (e.g., episodes into stories), through which

the sense of one's experience and a feeling of identity emerges (Mahoney & Granvold, 2005). The aim of narrative therapy is thus to "assist clients in revising their old stories" (Rosen, 1996, p. 24).

Narrative therapy addresses the criticism aimed against narrative medicine by designing research methodologies and providing empirical data on its effectiveness (see, e.g., Goncalves & Machado, 1999; Pennebaker, 1995). For example, Baerger & McAdams (1999) have shown that *coherence* of one's personal narrative, understood as saturation with the social context, the presence of an episodic structure (beginning, middle, end), and an emotional evaluation of the described events, is correlated with psychological wellbeing scores.

Postmodernism

The debates between various schools of psychotherapy take place in the wider context of *postmodernism*, in which increasing globalization, consumption, and awareness of the role language and political power play in shaping the dominant narrations is assumed to erode universal values (Hermans & Hermans-Konopka, 2010), including the positivist faith in empirical data which lies at the foundation of EBP.

This creates an obvious challenge for practitioners wanting to continue their education (see, e.g., Aarons, 2004; Hughes, 2004), but also for their clients, forced to create identities and find meanings in an increasingly complex, unstable, and idiosyncratic world in which "the pervasiveness of a global postmodern consumer culture [means] that most of us are faced with a confusing range of choices around lifestyle, occupation, religious and sexual orientation, and moral decision-making" (Angus & McLeod, 2004, p. 77). Such multiplicity can lead to oversaturation and blurring of identity (Gergen, 1991) on the one hand, and to emptiness on the other (Cushman, 1990). The challenge for contemporary narrations thus lies in accepting the coexistence of multiple possible perspectives. Psychotherapy can be a source of help in this regard (McLeod, 2004), particularly psychotherapy informed by *dialogical self theory* (DST), whose core assumption, interesting from the point of view of the postmodern reality and its challenges, is such a multiplicity of perspectives (Hermans & Hermans-Konopka, 2010).

Dialogical self theory

According to DST, the human personality should not be treated as a monolithic entity, but must rather reflect the socio-cultural conditions in which it is founded (Hermans & Gieser, 2012). Combining the notions of the *self*, which represents the dimension of space, and of *dialogue*, representing time, DST views personality as an *internal dialogical space* (Hermans, 2003) in which the self moves between various *I-positions* – parts or voices of the self, as well as the internalized representations of significant others – in the process of a dialogue, which shapes the personal narration (Hermans, 2014). The world of social roles and relationships thus becomes a part of one's identity, in turn influencing the way one acts upon this world (Hermans & Dimaggio, 2004). Moreover, these internal dialogues are assumed to mirror the rules governing external interpersonal relationships (Hermans & Kempen, 1993; Stemplewska-Żakowicz, Zalewski, Suszek, Kobylińska, & Szymczyk, 2012), so that some voices can dominate, others might be excluded, and others still – be unconscious and unrecognized (Hermans & Hermans-Konopka, 2010).

The dynamic and multifaceted nature of DST is also visible in its *bridging* (Hermans & Gieser, 2012) quality as a theory, combining many different constructs and methodologies. Within psychotherapy, the assumption of the mutual influence between *self*, *I*, and *mine* integrates shared concepts from such therapy approaches as humanistic, narrative, cognitive-relational or process-oriented (Hermans & Dimaggio, 2004). It thus seems to offer a useful framework for problems specific to postmodern narrations – problems that might be difficult to solve by educating clients into a single, dominant narration, such as the cognitive-behavioral "advocate of reality."

Dialogical self theory in psychotherapy

In DST-based therapy, clients' issues are treated as specific disturbances in their internal dialogues (Dimaggio, 2012). Their conceptualization is thus not reduced to dysfunctional elements. These are understood as parts of a wider repertoire of internal voices, and therapy offers a way of its restructuring (Hermans & Dimaggio, 2004) through correcting dialogical inequalities

– impoverishment, barrenness, repetitiveness, closing off to innovation, or disorganization (Dimaggio, 2012, p. 358). This can be done through increasing awareness of one's internal complexity and allowing new or hidden voices to come to the fore (Goncalves, Matos, & Santos, 2009). If the modern environment is complex and requires efficient participation in many (often contradictory) relations and roles, and if, according to DST, one's personality simultaneously reflects that environment and shapes it through influencing one's decision-making (Hermans & Dimaggio, 2004), then a well-functioning person should exhibit a wide array of internal voices (Dimaggio & Stiles, 2007).

The therapist not only analyzes the client's narration, but also gives them space to voice it within a safe relationship. Several measures – quantitative questionnaires (see, e.g., Hermans, 2001; Oleś, 2012) and projective tasks (Bokus, Bartczak, Szymańska, Chronowska, & Ważyńska, 2017) – have been designed within the DST framework, but the situation of the dialogue between the client and the therapist remains the central element (Opoczyńska, 2012), as "before the story reaches a final ending, it has taken another direction or turn under the influence of voices co-constructing new or alternative stories as part of the polyphony of the mind" (Hermans & Dimaggio, 2004, p. 2). The dialogue thus involves the client's voices, which they try to make sense of and communicate to the therapist, as well as the therapist's presence and voices responding to the client. The dialogue is situated in both space and time, and can be considered from a structural and a dynamic point of view (Chmielnicka-Kuter, 2012).

Such an approach can be especially useful with problems that fall outside of the so-called Axis I symptoms and instead refer to one's identity, sense of selfhood, or the continuity thereof. Examples include immigration, sudden and drastic changes in material conditions, a diagnosis of chronic illness, or the end of a long relationship (Chmielnicka-Kuter, 2012).

What can DST-informed therapy look like in practice? Traditional cognitive-behavioral therapy (CBT) sees depressive symptoms, such as lack of energy and motivation, less engagement in activities, and lowered mood as resulting from distorted, highly self-critical beliefs about oneself (e.g., "I am flawed") developed in childhood (Dobson & Dobson, 2018). Within DST, however, such symptoms can be described as effects of an impoverished internal dialogue: Few internal voices operating on a narrow

range of content, a lack of alternative viewpoints, or domination by especially critical significant other voices can lead to an equally distorted self-narration (Dimaggio, 2012).

The CBT therapist's work will be characterized by *collaborative empiricism* (Dobson & Dobson, 2018) through which the client is supported in developing a rational self-perspective (the therapist as the "advocate of reality"). An important source of this support is *Socratic dialogue* between the therapist and the client, also called *guided discovery* (Dobson & Dobson, 2018). This dialogue bears characteristics different than DST dialogues: By asking reflection-provoking questions ("What evidence do you have for this?" "Is it always this way?"), the therapist leads the patient towards a more objective outlook on their thoughts (e.g., "I have both strong and weaker sides"). In turn, this outlook allows for a more effective confrontation with various problems and an engagement in desired activities, also planned together with the therapist through so-called behavioral experiments (Bennett-Levy et al., 2004). The resulting experiences of agency will further reinforce a healthy perspective. In contrast, in DST-based therapy the therapist invites the client to voice their personal narration and fosters the identification of less pronounced voices and emotional states not made conscious (sequential reflecting of individual voices in a multivoiced story, Dimaggio & Stiles, 2007; using emotional expression techniques such as the "empty chair," Greenberg, 2002). The therapist has a chance of becoming a new voice in the client's repertoire (Hermans & Dimaggio, 2004). Such an internalized relationship can then contribute towards more effective emotional regulation and understanding of others' perspectives: "In their narrative landscape the [clients] give the therapist the role of a figure with whom they are in dialogue and discuss the meaning of the events that make them suffer" (Semerari, Carcione, Dimaggio, Nicolo, & Procacci, 2004, p. 232). Greater internal dialogical activity becomes the foundation of healthy self-esteem and interpersonal functioning (Dimaggio, 2012, see also Hermans, 2001, 2004; Whelton & Greenberg, 2004). Importantly, integration or the solving of all contradictions is not the goal of these dialogues – DST sees change and acceptance of multiplicity as sources of health (Hermans & Hermans-Konopka, 2010).

Though CBT and DST take distortions in the meaning-making process to be the cause of psychological problems, CBT focuses

on information processing and coping skills (e.g., social skills, Hope, Heimberg, & Turk, 2010), while DST emphasizes insight, internal representations of significant others, and the process of narrative identity shaping. CBT involves educating the client into the therapist's perspective, while DST-based therapy fosters the development of the client's internal dialogues, giving them greater freedom over what shape they will take. These differences become especially important when working with clients suffering from personality disorders. Traditional CBT, based on a rational analysis of cognitive content, can be perceived as invalidating and authoritative by such clients (Young, Klosko, & Weishaar, 2003). This creates a risk of declarative obedience out of fear of disappointing the therapist or of dropping out altogether (Semerari et al., 2004). Oriented towards giving clients space for working on their personal narrations, DST-based therapy can thus be more closely responsive to their individual needs and capabilities (Dimaggio & Stiles, 2007).

This can be illustrated with an example of working with clients suffering from paranoid personality disorder (Dimaggio, Catania, Salvatore, Carcione, & Nicolo, 2006; Dimaggio, 2012; Semerari et al., 2007). Within the DST framework, this disorder involves the narration of a self feeling weaker and flawed, shielding itself from perceived abuse by taking a mistrustful position (Salvatore, Nicolo, & Dimaggio, 2005). This narration is supported by an impaired ability to assume other perspectives (Semerari et al., 2007) and a vicious circle in which mistrust provokes negative reactions from others, confirming the initial mistrust. By treating the client's functioning as a constant process of meaning-making and dialogue between voices representing various experiences (weakness, fear, helplessness, Dimaggio, Semerari, Carcione, Nicolo, & Procacci, 2007), therapy focuses not on the content of paranoid cognitions, but on strong, unpleasant emotions that accompany the feeling of weakness and threat. Expressing these emotions strengthens new, previously disowned voices, which improves interpersonal functioning and self-esteem (Dimaggio et al., 2003).

DST-based therapy can also offer a productive framework for grief. Prolonged grief (over six months) has been distinguished as a diagnostic entity, separate from depression and post-traumatic stress disorder, in the DSM-V (Maciejewski, Maercker, Boelen, & Prigerson, 2016). Cognitive-behavioral programs have been

shown to be relatively effective, though less so than for many other disorders (see, e.g., Mancini, Griffin, & Bonanno, 2012). One reason might be the CBT conceptualization of grief. Aside from an incomplete integration of the experienced loss into autobiographical memory (experiencing loss as particularly salient), it underscores the role of behavioral avoidance of negative emotions and loss-related situations as well as negative beliefs about one's capacity to cope, and grief symptoms themselves (see Boelen, van den Hout, & van den Bout, 2006).

Based on this conceptualization, Rosner, Pfoh, and Kotoucova (2011) have proposed a 20- to 25-session CBT program. Its core involves educating the client about the cognitive characteristics of prolonged grief and restructuring loss-related cognitions. Clients are encouraged to consider such questions as "Is it realistic?" "Is it unrealistic?" "Is it realistic, but not helpful for me to think this way and do I have alternatives?" (p. 6). It is worth noting, however, that "special care is taken by the therapist not to reduce the suffering of anyone to the level of a mere dysfunctional thought, as otherwise the patients might feel invalidated" (p. 6). What is also important is that the program incorporates imaginal and emotion-focused techniques, such as leading an imaginary dialogue with the dead person. Other CBT programs for prolonged grief also include similar elements – repeated narrations about the loss (Boelen, de Keijser, van den Hout, & van den Bout, 2007), imaginary dialogues or therapeutic writing exercises (Jordan, Lubin, Larson, Wortmann, & Litz, 2016) – which points towards the potential usefulness of a narrative, dialogical approach to this issue.

"From a constructivist standpoint, grieving for the death of a loved one entails reaffirming or reconstructing a world of meaning that has been challenged by loss" (Neimeyer, 2012, p. 375). Sudden and difficult events, such as loss, are the greatest challenges for the continuity and coherence of personal narratives (Neimeyer & Arvay, 2004). Therapy centered around nurturing internal dialogues between voices of various perspectives on loss and the internalized representation of the dead person can thus help work through the loss (Neimeyer, 2006). This process engages the aforementioned experiential techniques to a greater extent. Clients are invited not to consider their own loss-related cognitions, but to voice contradictory or shameful emotions: "Far from being a cerebral, intellectual process, the quest for

meaning in both the loss and in their changed lives is typically a profoundly emotional one for survivors" (Neimeyer, 2012, p. 386). Externalizing and developing dialogues between parts of the self ("I as son/daughter," "I as a child," "I as strong/independent") and the dead person allows them to close the grieving process by integrating meaning. This constitutes a perspective complimentary to that of behavioral activation and pragmatic analysis of the functionality of one's cognition.

For and against

The above outline presents DST as complementary to the dominant CBT paradigm. What are the arguments for its use in practice, however?

Above all, the metaphor of personality as composed of multiple sides or voices seems to be accepted in both popular and scientific discourse (though see Suszek, 2012). Moreover, emphasizing flux and multiplicity without final integration (Hermans & Hermans-Konopka, 2010), DST responds to the problems of postmodern identity (Hermans, 2014) – healthy functioning means flexible adaptation to rapid changes. Also, identity is not only a monolithic entity based on universal values, but a process of meaning emerging from relationships – with significant others, with those values, with ideas, social group roles, and even with ourselves (Hermans & Hermans-Konopka, 2010). An approach to personality as a narration can thus be equally functional to shaping a rational self-perspective in CBT. Finally, discovering one's internal multiplicity and experimenting with various forms and expressions thereof can be a liberating experience for many, as it lifts from them, at least to some extent, the responsibility of seeking ideal unity and stability, even in the face of a changing reality (Dobson & Dobson, 2018): "Looking at one's life history as – partially – someone else's story allows one to gain distance to it and realize whether, and to what degree, one should continue it in the same way" (Chmielnicka-Kuter, 2012, p. 387).

On the other hand, DST is notable for its metaphorical character. The notion of psychological activity as a dialogue, described using such terms as *internal dialogical space*, *depositioning*, *shadow positions*, or *coalitions* (Hermans & Hermans-Konopka, 2010) is coherent and logically derived from interdisciplinary literature.

Nevertheless, to what extent will therapy based on such a metaphor accurately reflect the client's reality, and to what extent will it merely be a creative exercise in role-playing (Suszek, 2012)?

The answer might be found in another question: Does the responsibility for effective transfer of theory into practice lie with the theory, or the practitioner? Forcibly applying the internal dialogue and narration metaphor can be incompetent and can be met with resistance from the client, but this probability hardly seems unique to DST-based therapy. Most likely any therapist, regardless of their background, can at times feel uncertain in their work, and increased directiveness seems like an intuitive reaction. Indeed, a common criticism against CBT is that it relies on an overintellectualized register and prioritizes following the treatment manual over responding to the client's needs (Dobson & Dobson, 2018). Using the dialogue or narration metaphor with a client does not need to equal applying the DST framework in its entirety – such a false dichotomy is as unfair as treating CBT as purely "intellectual," consisting of nothing more than analyzing the client's automatic thoughts.

Next, using DST in therapy is not without precedent. The concept of a multifaceted self is "not new and has emerged several times in the course of the history of psychotherapy" (Suszek, 2012, p. 418). For example, *schema therapy* (ST, Young et al., 2003), lately gaining popularity and an increasing evidence base, is an extension of traditional CBT by aspects of meaning-making of relational and emotional experience – so-called schemata. Shaped in childhood and consisting of cognitions, memories, and emotions related to significant others (Young et al., 2003, p. 7), dysfunctional schemata resemble the notion of I-positions or internal voices which contain "specific cognitive-behavioral resources, formed by different ways of assigning meaning to personal experience" and which stem from "relationships with significant others, important social groups, or environmental influence" (Stemplewska-Żakowicz et al., 2012, p. 177). Schema therapy also distinguishes *modes*, or "moment-to-moment emotional states and coping responses – adaptive or maladaptive – that we all experience" (Young et al., p. 37). Modes are described in an even more personified way, having names such as the Happy Child or the Detached Protector. The similarities between some aspects of DST and ST thus constitute an argument for the productivity of the metaphor of personality as a process of interaction between parts.

At this point, however, another question emerges. Why, despite its relative attractiveness and potential usefulness (see Hermans & Gieser, 2012), does DST remain a niche theory? It seems that its novel and simultaneously metaphorical character cannot be the only contributing factor. Above all, DST is the area of interest for a narrow group of academic psychologists, and publications on DST largely concern its theoretical aspects or proposed areas of application. Valuable empirical data proving some parts of DST as well as diagnostic and research tools designed with internal dialogues in mind do exist (see e.g., Bokus et al., 2012; Oleś, 2012), but as regards psychotherapy, no single treatment manual nor any systematic examination of its effectiveness as a treatment modality has been published. The space for a constructive dialogue about DST is thus naturally limited and one can ask how to widen it – by publishing more evidence? Evidence of a different kind (e.g., RCT)? By popularizing the theory itself?

Taking into consideration the above arguments, is there a space for DST among the various schools of psychotherapy? As many voices coexist in the internal dialogical space, the coexistence of many therapeutic approaches – many ways of reaching a common goal – seems desirable. By responding to clients' needs, such pluralism fosters the field's healthy development. It is true that regardless of modality, psychotherapy has a shared aim: improving the client's functioning. It is also true that it achieves this aim largely through common factors (Laska, Gurman, & Wampold, 2014). Thus, perhaps it would be worthwhile to take a multivoiced stance and remember that truth can be reached in various ways.

References

Aarons, G. A. (2004). Mental health provider attitudes toward adoption of evidence-based practice: The Evidence-Based Practice Attitude Scale (EBPAS). *Mental Health Services Research*, 6(2), 61–74.

American Psychological Association (2013). *Recognition of psychotherapy effectiveness*. Retrieved from https://www.apa.org/about/policy/resolution-psychotherapy

Angus, L., McLeod, J. (2004). Self-multiplicity and narrative expression in psychotherapy. In: H. J. M. Hermans, G. Dimaggio (Eds.), *The Dialogical Self in Psychotherapy* (pp. 77–91). Hove: Taylor & Francis.

Baerger, D. R., McAdams, D. P. (1999). Life story coherence and its relation to psychological well-being. *Narrative Inquiry, 9(1), 69–96.*

Bennett-Levy, J. E., Butler, G. E., Fennell, M. E., Hackman, A. E., Mueller, M. E., & Westbrook, D. E. (2004). *Oxford Guide to Behavioural Experiments in Cognitive Therapy*. Oxford: Oxford University Press.

Boelen, P. A., de Keijser, J., van den Hout, M. A., & van den Bout, J. (2007). Treatment of complicated grief: A comparison between cognitive-behavioral therapy and supportive counseling. *Journal of Consulting and Clinical Psychology, 75,* 277–284.

Boelen, P. A., van den Hout, M. A., & van den Bout, J. (2006). A cognitive-behavioral conceptualization of complicated grief. *Clinical Psychology: Science and Practice*, 13(2), 109–128.

Bokus, B., Bartczak, M., Szymańska, A., Chronowska, R., & Ważyńska, A. (2017). The dialogical self's round table: Who sits at it and where? *Psychology of Language and Communication*, 21(1), 84–108.

Charon, R. (2001). Narrative medicine: A model for empathy, reflection, profession, and trust. *The Journal of the American Medical Association*, 286, 1897–1902.

Charon, R. (2012). At the membranes of care: Stories in narrative medicine. *Academic Medicine*, 87(3), 342–347.

Chmielnicka-Kuter, E. (2012). W dialogu z własną historią – metafora (wielogłosowej) powieści w poradnictwie psychologicznym [In dialogue with one's own history: The metaphor of a (multivoiced) novel in psychological counseling]. In: P. Oleś, M. Puchalska-Wasyl (Eds.), *Dialog z samym sobą* [Dialogue with Oneself] (pp. 376–304). Warszawa: Wydawnictwo Naukowe PWN.

Cushman, P. (1990). Why the self is empty: Toward a historically situated psychology. *American Psychologist*, 45(5), 599–611.

David, D., Montgomery, G. H. (2011). The scientific status of psychotherapies: A new evaluative framework for evidence-based psychosocial interventions. *Clinical Psychology: Science and Practice*, 18(2), 89–99.

Dimaggio, G. (2012). Dialogically oriented therapies and the role of poor metacognition in personality disorders. In: H. J. M. Hermans, T. Gieser (Eds.), *Handbook of Dialogical Self Theory* (pp. 356–373). Cambridge, MA: Cambridge University Press.

Dimaggio, G., Catania, D., Salvatore, G., Carcione, A., & Nicolò, G. (2006). Psychotherapy of paranoid personality disorder from the perspective of dialogical self theory. *Counselling Psychology Quarterly*, 19(01), 69–87.

Dimaggio, G., Salvatore, G., Azzara, C., Catania, D., Semerari, A., & Hermans, H.J.M. (2003). Dialogical relationship in impoverished narratives: From theory to clinical practice. *Psychology and Psychotherapy, 76,* 4, 385–410.

Dimaggio, G., Semerari, A., Carcione, A., Nicolo, G., Procacci, M. (2007). *Psychotherapy of Personality Disorders: Metacognition, States of Mind and Interpersonal Cycles*. London: Routledge.

Dimaggio, G., Stiles, W. B. (2007). Psychotherapy in light of internal multiplicity. *Journal of Clinical Psychology*, 63(2), 119–127.

Dobson, D., Dobson, K. S. (2018). *Evidence-Based Practice of Cognitive--Behavioral Therapy*. New York: Guilford Press.

Garfield, S. L. (2006). Therapies-Modern and popular: PsycCRITIQUES 2006. Washington, DC: American Psychological Association.

Gergen, K. J. (1991). *The Saturated Self: Dilemmas of Identity in Contemporary Life*. London: Sage.

Ghaemi, S. N. (2007). *The Concepts of Psychiatry: A Pluralistic Approach to the Mind and Mental Illness*. Baltimore: JHU Press.

Gonçalves, Ó. F., Machado, P. P. (1999). Cognitive narrative psychotherapy: Research foundations. *Journal of Clinical Psychology*, *55*(10), 1179–1191.

Goncalves, M., Matos, M., & Santos, A. (2009) Narrative therapy and the nature of 'innovative moments' in the construction of change. *Journal of Constructivist Psychology, 22*, 1–23

Goyal, R. K., Charon, R., Lekas, H. M., Fullilove, M. T., Devlin, M. J., Falzon, L., & Wyer, P. C. (2008). 'A local habitation and a name': How narrative evidence-based medicine transforms the translational research paradigm. *Journal of Evaluation in Clinical Practice*, *14*, 732–741.

Grant, A. (2009). Evidence-based practice and the need for paradigmatic pluralism in cognitive behavioural psychotherapy. *Journal of Psychiatric and Mental Health Nursing*, *16*(4), 368–375.

Greenberg, L. S. (2002). *Emotion-Focused Therapy: Coaching Clients to Work Through Their Feelings*. Washington: APA Press.

Hayes, S. C., Hofmann, S. G. (2018). *Process-Based CBT: The Science and Core Clinical Competencies of Cognitive Behavioral Therapy*. Oakland: New Harbinger Publications.

Hermans, H. J. M. (2001). The dialogical self: Toward a theory of personal and cultural positioning. *Culture and Psychology, 7*, 243–248.

Hermans, H. J. M. (2003). The construction and reconstruction of a dialogical self. *Journal of Constructivist Psychology*, *16*(2), 89–130.

Hermans, H. J. M. (2004). The innovation of self-narratives: A dialogical approach. In: L. Angus, & J. McLeod (Eds.), *Handbook of Narrative and Psychotherapy: Practice, Theory and Research* (pp. 175–192). Thousand Oaks: Sage.

Hermans, H. J. (2014). Self as a society of I-positions: A dialogical approach to counseling. *The Journal of Humanistic Counseling*, *53*(2), 134–159.

Hermans, H. J. M., Dimaggio, G. (2004). The dialogical self in psychotherapy: Introduction. In: H. J. M. Hermans, G. Dimaggio (Eds.), *The Dialogical Self in Psychotherapy* (pp. 1–10). Hove: Taylor & Francis.

Hermans, H. J., Gieser, T. (Eds.). (2012). *Handbook of Dialogical Self Theory*. Cambridge, MA: Cambridge University Press.

Hermans, H., Hermans-Konopka, A. (2010). *Dialogical Self Theory: Positioning and Counter-Positioning in a Globalizing Society*. Cambridge, MA: Cambridge University Press.

Hermans, H. K. M., Kempen, H. J. G. (1993). *The Dialogical Self: Meaning as Movement*. San Diego: Academic Press.

Holmes, D., Murray, S. J., Perron, A., & Rail, G. (2006). Deconstructing the evidence-based discourse in health sciences: Truth, power and fascism. *International Journal of Evidence-Based Healthcare*, *4*(3), 180–186.

Hope, D. A., Heimberg, R. G., & Turk, C. L. (2010). *Managing Social Anxiety: A Cognitive-Behavioral Therapy Approach: Therapist Guide*. New York: Oxford University Press.

Hughes, B. M. (2008). How should clinical psychologists approach complementary and alternative medicine? Empirical, epistemological, and ethical considerations. *Clinical Psychology Review*, *28*(4), 657–675.

Jordan, A., Lubin, R., Larson, R., Wortmann, J., & Litz, B. (2016). Treating prolonged grief. In: H. Friedman, *Encyclopedia of Mental Health* (pp. 341–344). New York: Elsevier.

Koerner, K. (2018). Science in practice. In: S. C. Hayes, S. G. Hofmann (Eds.), *Process-Based CBT: The Science and Core Clinical Competencies of Cognitive Behavioral Therapy* (pp. 45–65). Oakland: New Harbinger Publications.

Kozlowski, D., Hutchinson, M., Hurley, J., Rowley, J., & Sutherland, J. (2017). The role of emotion in clinical decision making: An integrative literature review. *BMC Medical Education*, *17*(1), 255.

Laska, K. M., Gurman, A. S., & Wampold, B. E. (2014). Expanding the lens of evidence-based practice in psychotherapy: A common factors perspective. *Psychotherapy, 51*(4), 467–481.

Machado, P. P., Gonçalves, Ó. F. (1999). Introduction: Narrative in psychotherapy: The emerging metaphor. *Journal of Clinical Psychology*, *55*(10), 1175–1177.

Maciejewski, P. K., Maercker, A., Boelen, P. A., & Prigerson, H. G. (2016). "Prolonged grief disorder" and "persistent complex bereavement disorder", but not "complicated grief", are one and the same diagnostic entity: An analysis of data from the Yale Bereavement Study. *World Psychiatry*, *15*(3), 266–275.

Mahoney, M. J., Granvold, D. K. (2005). Constructivism and psychotherapy. *World Psychiatry*, *4*(2), 74–77.

Mancini, A. D., Griffin, P., & Bonanno, G. A. (2012). Recent trends in the treatment of prolonged grief. *Current Opinion in Psychiatry*, *25*(1), 46–51.

McLeod, J. (2004). Social construction, narrative and psychotherapy. In: L. Angus, J. McLeod (Eds.), *Handbook of Narrative and Psychotherapy: Practice, Theory and Research* (pp. 351–366). Thousand Oaks: Sage.

Neimeyer, R. A., Arvay, M. J. (2004). Performing the self: Therapeutic enactment and the narrative integration of loss. In: H. J. M. Hermans, G. Dimaggio (Eds.), *The Dialogical Self in Psychotherapy* (pp. 173–189) Hove: Taylor & Francis.

Neimeyer, R. A. (2006). Narrating the dialogical self: Toward an expanded toolbox for the counselling psychologist. *Counselling Psychology Quarterly, 19*, 105–120.

Neimeyer, R. A. (2012). Reconstructing the self in the wake of loss: A dialogical contribution. In: H. Hermans, T. Gieser (Eds.), *Handbook of Dialogical Self Theory* (pp. 374–389). Cambridge: Cambridge University Press.

Oleś, P. K. (2012). Dialogowe Ja: zarys teorii, inspiracje badawcze, ciekawsze wyniki [The dialogical self: Outline of the theory, research inspirations, interesting results]. In: P. K. Oleś, M. Puchalska-Wasyl (Eds.), *Dialog z samym sobą* [Dialogue with Oneself] (pp. 143–175). Warszawa: Wydawnictwo Naukowe PWN.

O'Mahony, S. (2013). Against narrative medicine. *Perspectives in Biology and Medicine*, *56*(4), 611–619.

Opoczyńska, M. (2012). Słowem, z powrotem do rzeczy. O dochodzeniu do siebie w psychoterapii [In a word, back to matters: On getting your self back in psychotherapy]. In: P. Oleś, M. Puchalska-Wasyl (Eds.), *Dialog z samym sobą* [Dialogue with Oneself] (pp. 360–376). Warszawa: Wydawnictwo Naukowe PWN.

Orlinsky, D. (2017). Unity and diversity among psychotherapies. In: A. J. Consoli, L. E. Beutler, & B. Bongar (Eds.), *Comprehensive Textbook of Psychotherapy: Theory and Practice* (pp. 11–31). New York: Oxford University Press.

Paul, G. L. (1969). Behavior modification research: Design and tactics. In: C. M. Franks (Ed.), *Behavior Therapy: Appraisal and Status* (pp. 29–62). New York: McGraw-Hill.

Pennebaker, J. W. (1995). *Emotion, Disclosure and Health*. Washington: APA Press.

Rosen, H. (1996). Meaning-making narratives: Foundations for constructivist and social constructionist psychotherapies. In: H. Rosen, K. Kuehlwein (Eds.), *Constructing Realities* (pp. 3–54). San Francisco: Jossey-Bass.

Rosner, R., Pfoh, G., & Kotoučová, M. (2011). Treatment of complicated grief. *European Journal of Psychotraumatology*, 2(1), 7995.

Salvatore, G., Nicolò, G., & Dimaggio, G. (2005). Impoverished dialogical relationship patterns in paranoid personality disorder. *American Journal of Psychotherapy*, *59*(3), 247–265.

Schultz, D. P., Schultz, S. E. (2008). *Historia współczesnej psychologii* [History of Contemporary Psychology]. Kraków: Wydawnictwo Uniwersytetu Jagiellońskiego.

Semerari, A., Carcione, A., Dimaggio, G., Nicolo, G., & Procacci, M. (2004). A dialogical approach to patients with personality disorders. In: H. J. M. Hermans, G. Dimaggio (Eds.), *The Dialogical Self in Psychotherapy* (pp. 220–234). Hove: Taylor & Francis.

Showalter, E. (1985). *The Female Malady: Women, Madness and English Culture, 1830–1980*. London: Virago.

Stemplewska-Żakowicz, K., Zalewski, B., Suszek, H., Kobylińska, D., & Szymczyk, B. (2012). Architektura poznawcza dialogowego Ja: umysł dyskursywny [Cognitive architecture of the dialogical self: The discursive mind]. In: P. K. Oleś, M. Puchalska-Wasyl (Eds.), *Dialog z samym sobą* [Dialogue with Oneself] (pp. 175–200). Warszawa: Wydawnictwo Naukowe PWN.

Suszek, H. (2012). Krytyka teorii dialogowego Ja [Critique of dialogical self theory]. In: P. Oleś, M. Puchalska-Wasyl (Eds.), *Dialog z samym sobą* [Dialogue with Oneself] (pp. 394–430). Warszawa: Wydawnictwo Naukowe PWN.

Whelton, W. J., Greenberg, L. S. (2004). From discord to dialogue: Internal voices and the reorganization of the self in process-experiential therapy. In: H. J. M. Hermans, G. Dimaggio (Eds.), *The Dialogical Self in Psychotherapy* (pp. 124–139). Hove: Taylor & Francis.

Young, J. E., Klosko, J. S., & Weishaar, M. E. (2003). *Schema Therapy: A Practitioner's Guide*. New York: Guilford Press.

Anna Milanowicz
ORCID 0000-0003-2833-9084
Faculty of Psychology, University of Warsaw

Truth and Untruth in Irony

Abstract

The concept of truth in irony should be approached through the lens of the speaker's beliefs and intentions. However, the experience and emotions of an ironic message's recipient are not without significance. The proposed approach describes irony as a form of concealment and camouflage. Irony vibrates between "what is said but not intended" and "what is intended but not said." It balances between the ostensible truth and the ostensible untruth. "Ironic criticism," also known as "blaming by praise," appears positive, although its true message is not. The reverse happens to be true for "ironic praise." From a moral perspective, irony is not virtuous: It violates "truth-telling" and is sometimes clownish. The author's final considerations on the subject of "Does irony have gender?" are followed by a concise description of gender differences in the use of irony and self-mockery.

Keywords: truth, non-literal language, irony, meaning ambiguity, gender

Using irony means using such "ways of speaking, writing, acting, behaving, painting, etc., in which the real or intended meaning presented or evoked is intentionally quite other than, and incompatible with, the ostensible or pretended meaning" (Muecke, 2002, p. 61)[1].

It is ironic that irony itself is often characterized as "pragmatic insincerity" while it is actually a sincere, if masked, evocation of the truth. The classic paradox of Socratic irony, "I know that I know nothing" (from Greek: Οἶδα οὐδὲν εἰδώς [Oída oudén eidós]), is as much true as it is false. Socratic irony is typically defined as the feigning of ignorance in challenging others. It also creates an unsettling tension between ignorance and knowledge: "You know I don't quite mean this, but what do I mean?" It is sometimes

---

[1] English text from Muecke, D.C. (1969). *The Compass of Irony*. London: Methuen & Co Ltd., p. 53.

considered that Socrates was expressing the futility of trying to obtain absolute knowledge. This ambiguity of irony between the known *versus* the unknown, joking *versus* criticizing, inclusion of those who understand the irony *versus* exclusion of those from outside the group: these were the reasons why Socrates was put on trial by the Athenians. Teetering on the brink of illusionary truth and illusionary untruth is both powerful and dangerous.

The nature of irony, immersed in the spirit of ancient Greece, might be a reflection of our passion for word play. **Dramatic irony** as it appears on the ancient stage is dressed in the cloak of an underdog, the *eirôn*, who uses irony to trick a boastful bully, the *alazṓn*. The *alazṓn* remains unaware of this technique, unlike the audience. Irony here is about power. This "inclusion" and initiation into irony must naturally come from the mystical tradition of ancient Greece. There is not only something "pretend" about irony but also something sublime on the way to real knowledge. Eventually the *eirôn* rises morally above his opponent. Even though the boastful *alazṓn* appears to dominate the stage, it is the *eirôn* who establishes rapport with the audience and wins them over.

Dramatic irony provides the very foundation of storytelling in literature and contemporary films, where the audience have an advantage over the characters because we are on different ends of the unfolding "drama." We see the truth of what is going on. Any stage directions, remarks or narrator's comments are additional perspectives that supplement each character's own perspective. The audience has a god-like perspective, while the characters are in the shadow of partial knowledge.

We criticize with irony more than we compliment with irony (Maciuszek, 2018). This might result from a social agreement between language users that obliges us to be nice rather than unpleasant in social interactions (Boucher & Osgood, 1969; Gibbs, 1986). "Ironic criticism," also known as "blaming by praise," appears positive, although its true message is not. This indirect criticism is used significantly more in language than "ironic compliments," or "praising by blame." Also, according to Rachel Giora (1995) detection of ironic compliments requires a double negation, which makes it more complicated to process than the single negation of ironic criticism.

The commonly accepted view is that negation in irony – being UNclear, leaving things UNsaid, saying UNtruths, being

INsincere – implies disappointment of expectations or norms (Sperber & Wilson, 1981; Kumon-Nakamura et al., 1995; Utsumi, 2001). Let us consider an alternative, i.e. that it instead implies absence of disappointment: that irony signals that the speaker's previous expectations have been disappointed and there is now a despondent lack of expectation. For example, let's take a situation when someone disappoints us yet again, and comments are exchanged: the recipient's question, "What do you expect?" is countered by the ironic speaker, "I have no expectations." The response reveals that there were once hopes, but that, having been let down before, we actually do not expect things to be any different this time; rather, we just wish this state of affairs were different. The colloquial "Great!" becomes ironic when things are not "great," but we did not expect them to be otherwise.

Similarly, the statement "You're so smart!" assumes an ironic tone when the person on the receiving end was not expected to act other than stupidly, probably not for the first time. We use irony not to express "failed expectations" but the absence of any positive expectations. We might call this Negation of Expectations in Irony (NEI). NEI also explains why irony is used more often in close relationships. Only in close relationships or repetitive situations[2] can one observe the regularity of certain behaviors and, consequently, develop a lack of expectation that things would be different. Irony then does not express feelings of disappointment or failed expectations.

Even the classic example of the ironic "great weather" implies the idea that the ironic speaker expected it to rain, and he or she is simply expressing what they think it would ideally be like. In verbal irony, similarly to the Socratic pursuit of "absolute knowledge," the desired state of affairs always remains illusory and unreachable. Both the ironic speaker and the "intimates" from within the ironic inclusion circle had no other expectations whatsoever.

Inclusion into "pragmatic insincerity" always comes with attention sharing. This type of joint attention is often possible because of the use of language. Therefore, INcomprehension of irony does not make the speaker a liar. In typical irony, the

---

[2] An ironist will be likely to use irony to comment on a state of affairs they know well, such as politics or the weather, where the fact that things have gone wrong again comes as no surprise.

speaker does not say what they think and what they think is not what they say. Example: "Great idea!" when they think it's actually bad. With respect to the cooperative principle of Grice (1975), according to which listeners and speakers act rationally and cooperate in order to achieve effective conversational communication, violation of the maxim of quality – *tell the truth* – results in a conversational *implicature* of the proper sense of an ironic comment, its meaningful truth. The cooperative principle is kept. However, there is another category of ironic comments: those based on explicit truthfulness.

In the case of verbal comments reflecting a true state of affairs or at least appearing to represent such truth (*verisimilar irony*, see Dynel, 2018), the speaker says something that does not require evaluation reversal, e.g. this classic example: "I like children who keep their rooms clean" said by a mother coming into her son's messy room (Gibbs & O'Brien, 1991; Hamamoto, 1998; Utsumi, 2000). The child who is the recipient of the verbal irony does not keep his room clean, at least not on this occasion. Here, "children who keep their rooms clean" are given as an example to follow. The mother who uses irony wants also her child to keep his room clean. Once again, verbal irony is the expression of some general, desired yet unobtainable ideal. Again, there is no disappointment here, since the "mess" was probably expected.

A linguistic expression has no absolute "meaning" outside of its specific context (Wittgenstein, 2005). This context includes both the objective reality of the situation and the subjective perceptions and thoughts of the speaker and recipients. Therefore, the truthfulness of a linguistic expression has to be considered through the lens of the speaker's intentions. Irony, in this respect being similar to lying, must be addressed to someone, i.e. the existence of an addressee is a necessary condition (Mahon, 2008). However, unlike lying, it is expected that the listener understands the intended meaning (Siegler, 1996). Both irony and lying can be used in self-defense and as face-saving techniques. One might become a victim of irony just like one becomes the victim of a lie. However, it is the intention of the speaker that differs.

A lie is an untruthful statement, expressed by the speaker with the intention of making someone else believe it as true (Primoratz, 1984). Thus, the intention of a liar is to trick the recipient into believing that what is untrue is true. Since the intention of a person using irony is its detection by the recipient,

if not an addressee then a listener[3], an ironic statement – even if it sounds untrue – is not a lie.

Irony as a manifestation of non-literal language balances between what is said but not thought and what is thought but not said. From a moral perspective, irony is not virtuous: It violates "truth-telling" and is sometimes clownish. It is a form of masking and camouflage. Irony is generated from the speaker's true opinions and feelings, while he or she is assuming or pretending something else (see the pretense theory of Clark & Gerrig, 1984).

Does irony have gender?

Our research (Milanowicz, Tarnowski, Bokus, 2017) shows that although verbal "irony" is etymologically female (Polish: *ironia* – feminine grammatical gender), its daily application is more of a male domain. Irony is used more frequently by men than women. Men also generate more irony in response to irony, in the case of both "ironic criticism" (blame by praise) and "ironic praise" (praise by blame).

The important thing is, however, WHO is speaking to WHOM. Women tune into ironic mode more with men than other women. Men, however, use less irony towards women than towards men. Also, higher self-esteem and lower anxiety seem to be good predictors of the use of self-mockery. Women seem to use more self-deprecation, whereas men are more likely to use irony as positive self-affirmation (Milanowicz, Bokus, 2020).

Men perceive speakers' intentions as more positive than women. More women than men believe that an ironist's intentions were driven by meanness and hostility. Men also experience more positive emotions as recipients of ironic comments. Men feel more joy, while the dominating feelings among women are anger and sadness (Milanowicz, 2019).

In social psychology, the duality of gender stereotypes is expressed in the dimensions of "agency" and "communion": women regard themselves as more communal while men regard themselves as more agentic (Wojciszke, 2010). It is possible that the male proclivity for "ironic ambiguity" reflects this dimension of "agency" in language. Using non-literal language becomes a manifestation of power and control not only over the message but also its recipients.

---

[3] Verbal irony, like dramatic irony, can be addressed to an interlocutor who is its "victim," but it is the listener, observer or audience that is actually meant to understand it.

References

Boucher, J., Osgood, C. E. (1969). The Pollyanna hypothesis. *Journal of Verbal Learning and Verbal Behavior*, *8*(1), 1–8.

Clark, H., Gerrig, R. (1984). On the pretense theory of irony. *Journal of Experimental Psychology: General*, *113*, 121–126. doi: 10.1037/0096-3445.113.1.121

Dynel, M. (2018). *Irony, Deception and Humour: Seeking the Truth about Overt and Covert Untruthfulness*. Boston: Mouton de Gruyter.

Gibbs, R. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General*, *115*(1), 3–15.

Gibbs, R.W., O'Brien, J. (1991). Psychological aspects of irony understanding. *Journal of Pragmatics, 16,* 523–530.

Giora, R. (1995). On irony and negation. *Discourse Processes*, *19*, 239–264. doi: 10.1080/01638539509544916

Grice, H. (1975). Logic and conversation. In: P. Cole, J. L. Morgan (Eds.), *Syntax and Semantics*, *3: Speech Acts* (pp. 41–58). New York: Academic Press.

Hamamoto, H. (1998). Irony from a cognitive perspective. In: R. Carston, S. Uchida (Eds.), *Relevance Theory, Applications and Implications* (pp. 257–270). Amsterdam: John Benjamins.

Kumon-Nakamura, S., Glucksberg, S., & Brown, M. (1995). How about another piece of pie: The allusional pretense theory of discourse irony. *Journal of Experimental Psychology: General*, *124*(1), 3–21. doi: 10.1037/0096-3445.124.1.3

Maciuszek, J. (2018). Jak kontekst wpływa na znaczenie przekazu; o rozumieniu pośrednich i niedosłownych wypowiedzi (metafory, ironia, pośrednie akty mowy, wieloznaczności) [How context affects meaning; understanding indirect and non-literal statements (metaphor, irony, indirect speech acts, ambiguity)]. In: B. Bokus, E. Kosowska (Eds.), *Człowiek w sytuacji – nie tylko z perspektywy psychologa. Studia inspirowane teorią Tadeusza Tomaszewskiego* [Humans in Situations, not only from a Psychologist's Perspective. Studies Inspired by the Theory of Tadeusz Tomaszewski] (pp. 338–359). Warszawa: Wydawnictwa Uniwersytetu Warszawskiego & SWPS Uniwersytet Humanistycznospołeczny.

Mahon, J. E. (2008). Two definitions of lying. *International Journal of Applied Philosophy 22*(2), 211–230.

Milanowicz, A. (2019). Stereotyp płci w odniesieniu do języka niedosłownego. Z badań nad ironią. [Gender stereotypes related to non-literal language: The case of irony. Doctoral dissertation supervised by B. Bokus], University of Warsaw.

Milanowicz, A., Bokus, B. (2020). *W krzywym zwierciadle ironii i autoironii. O kobietach i mężczyznach nie wprost* [The Distorted Lens of Irony and Self-Mockery: A Gender Comparison]. Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Milanowicz, A., Tarnowski, A., & Bokus, B. (2017). When sugar-coated words taste dry: The relationship between gender, anxiety, and response to irony. *Frontiers in Psychology 8*, 2215. Front. Psychol., 19 December 2017 DOI: doi.org/10.3389/fpsyg.2017.02215

Muecke, D. C. (2002). Ironia: podstawowe klasyfikacje [Irony: Basic Classifications], transl. G. Cendrowska. In: M. Głowiński (Ed.), *Ironia* [Irony] (pp. 43–74). Gdańsk: słowo/obraz terytoria.

Primoratz, I. (1984). Lying and the "methods of ethics". *International Studies in Philosophy*, *16*, 35–57.

Siegler, R. S. (1996). *Emerging Minds: The Process of Change in Children's Thinking.* New York: Oxford University Press.

Sperber, D., Wilson, D. (1984). Pragmatics: An overview. In: S. Georges (Ed.), *From the Linguistic to the Social Context* (pp. 21–41). Bologna: Cooperativa Libraria Universitaria Editrice.

Utsumi, A. (2000). Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from non-irony. *Journal of Pragmatics*, *32*(12), 1777–1806. doi: 10.1016/S0378-2166(99)00116-2

Wittgenstein, L. (2005). *Dociekania filozoficzne* [Philosophical Inquiries]. Warszawa: Wydawnictwo Naukowe PWN.

Wojciszke, B. (2010). *Sprawczość i wspólnotowość. Podstawowe wymiary spostrzegania społecznego* [Agency and Community. Basic Dimensions of Social Perception]. Gdańsk: Gdańskie Wydawnictwo Psychologiczne.

Adrianna Smurzyńska
ORCID 0000-0001-9926-576X
Jagiellonian University
»Artes Liberales« Academy

When Does Simulation Enable Us Adequately to Attribute Mental States to Others?

Abstract

The paper discusses issues of taking different perspectives and attributing mental states. The main focus is on the criteria of their adequate ascription. Taking another's perspective and simulation theory are taken into consideration. The paper considers the results of research on the development of the ability to differentiate perspectives and disorders of this ability in borderline personality disorder. The ability to differentiate one's own and others' mental states has been presented as an aspect of social cognition that enables adequate mentalization, by accepting the presence of various perspectives and the possibility of the coexistence of various representations of reality.

Keywords: mentalization, simulation, differentiation of perspectives

1. Introduction

When we look at others, we make inferences about their way of thinking. We try to judge what their general beliefs are, what they feel when they are smiling or crying, what they are going to do when they perform certain actions. When we look at people gazing into each other's eyes, we conclude that they are in love. When a student studies conscientiously for exams, we not only assume that he or she cares about them but also that he or she is a hard-working person. When we see a crying child sitting on the sidewalk with its knee skinned, not only do we recognize that it feels pain, but also that it needs help and is wondering what to do. There are countless such examples. They indicate a widespread necessity for answering such questions as: What is he thinking? What does she feel? What do they think about

that? In philosophy, psychology, and other branches of cognitive science, the ascription of mental states is described by the term "mentalization."

The present paper discusses the issue of taking different perspectives and attributing mental states. I will focus mainly on the criteria of their adequate ascription. I will refer mostly to simulation theory and to research on mentalization in personality disorders.

2. What is mentalization?

The problem of other minds has been present in philosophy for a long time. In the past, it was stated that we have direct and private access to our own mind and that our beliefs about ourselves are infallible (cf. Hyslop, 2016). According to one of the most famous theories of the other minds problem, knowledge about our own mind was regarded as the basis for inferring knowledge about others (cf. Russell, 1970). It was assumed that if someone knows that his or her behavior X reflects feeling Y, then seeing such behavior of person B, he or she could tell that person B was feeling Y.

Nowadays, it is doubted whether we have private access to our own mental states. Beliefs about oneself are not assumed to be infallible anymore, either. Compared to the classical theories, in most contemporary theories there is no place for privileged and direct access to one's own mind. Moreover, the adequacy of mental state attribution to oneself becomes problematic. Such an approach is grounded in the results of famous social psychology experiments such as Milgram's experiment or Zimbardo's experiment. Beliefs about oneself are not assumed to be reliable and unquestionable but are thought to be fallible and biased.

The capability to attribute mental states is one of the main issues in research on social cognition, i.e. cognitive abilities which are specific for social interactions (Vogeley & Newen, 2009). This capability consists in explaining and predicting others' behavior and enables effective communication and interaction between people (Vogeley & Newen, 2009). It is described as the ability to represent beliefs about others' mental states (cf. Goldman, 2009), understanding others (Robbins, 2004), ascribing mental states to oneself and to others (Goldman, 2012).

Such a competence in literature is called theory of mind, mindreading or mentalization (cf. Vogeley & Newen, 2009). In this paper I use the third term: mentalization, because in my opinion it appears to describe the process in the most neutral way, without predetermining its mechanism (cf. Guerini & Marraffa, 2015; Robbins, 2004).

3. How do we mentalize?

The question of how we understand others has led to the creation of several theories attempting to describe this mechanism. Some research claims that we have a theory (innate or learned) of people's way of thinking and behaving (cf. Perner, 1991). Other bodies of research assert that we build person models of people's functioning (Newen, 2015). Another theory is that we have the ability to simulate others' mental states (Goldman, 2006) or we perceive them directly (Gallagher, 2008). Each of these conceptions refers to psychology research, but it is difficult to assume which of them describes this competence completely. It is probable that the ability to ascribe mental states is a heterogenous process and people use different epistemic strategies (e.g. those mentioned above), choosing one of them depending on the situation (cf. Newen, 2015).

In this paper I will invoke one of these conceptions, the one that focuses on perspective-taking ability: Simulation Theory. In this case, understanding others is assumed to be the result of simulating others' mental states (Goldman & Mason, 2007; Newen, 2015). The literature includes many different approaches to this type of process (Barlassina & Gordon, 2017). In the classical versions (e.g. Gordon, 1995), simulation was understood as an automatic and unconscious process. Unorthodox versions could distinguish low-level unconscious processes such as mirroring, and high-level processes such as reflection and imagination (Goldman, 2006; Waytz & Mitchell, 2008). In this case, the process of understanding others – when someone tries to put themself in another's shoes – happens both on an intuitive as well as on a reflexive level.

Our own mind is treated as a tool which enables us to take another person's perspective (Goldman, 2006; Newen, 2015; Waytz & Mitchell, 2008). In general, the simulation process consists in imagining another's mental state and attributing this

state to that person (Bogdan, 2007; Goldman, 2006; Krueger & Overgaard, 2012; Newen & Schlicht, 2009). Imagining another person's mental state is described as re-creating or re-enacting the other's mental states (Goldman, 2009). It is a kind of mental pretense of another person's mental life (Goldman, 2009). Some advocates of this conception suggest that people have the ability of introspection or that they make use of their self-understanding when understanding others (cf. Goldman, 2006; Mitchell, Macrae, & Banaji, 2006). Nevertheless, most supporters of simulation theory admit that understanding oneself is based on the same mechanisms as understanding others (cf. Waytz & Mitchell, 2008).

4. When does the simulation process lead to adequate mental state ascription?

Proponents of simulation theory claim it is possible to use one's own mind for understanding others. At the same time, they indicate that this process is different from projecting one's own feelings onto others (cf. Goldman, 2006). On the other hand, it has been scientifically proven that one's own experiences influence mental state recognition. Experiencing specific mental states enables one later to attribute them to others (Decety & Grezes, 2006). On the other hand, it is possible that when persons A and B have experienced a similar situation, person A may attribute his or her feelings to B or react to B less empathetically than someone who did not have such an experience (cf. Hodges, 2005).

Putting oneself in another's shoes happens through using one's own imaginings and experience as an anchor when trying to understand others (Waytz & Mitchell, 2008). Quite often one's own beliefs or emotions are inadequately attributed to others (Goldman & Mason, 2007). Problems with differentiating the two perspectives: one's own and another's, seem to be much more frequent than it was assumed (Hodges, 2005). The separation of two similar experiences requires special cognitive capabilities and demands intellectual effort (Boyer, Robbins, & Jack, 2005).

Therefore we need to answer the question in what kind of situations it is possible to ascribe mental states adequately. One proposal invokes similarities between the interacting persons (Ames, 2007; Frith & Frith, 2006; Newen, 2015). If person A has similar traits and experiences as person B, it is highly

probable that person A will properly attribute mental states to person B. In this case, it is necessary to warrant the statement that another person is similar or dissimilar with regard to the mental state being considered. This situation enables a person to use self-knowledge when explaining or predicting another's behavior (Newen, 2015). It can be estimated that adequate mental state ascription in this case is highly probable.

We need to ask, though, if in this case any new information is gained about the other person. If person A knows that person B is similar to him or her, he or she does not have to simulate B's mental states. In this case, it is enough to recognize the behavior and refer to the knowledge of oneself and person B. Adequate mental state ascription in this case would only be the result of memorizing previous meetings with person B. The question of how new information is obtained still remains open, and similarities seem to be insufficient for ascribing mental states adequately.

Some commentators of simulation theory indicate that for appropriate mental simulation, the abilities of mirroring and of differentiating oneself and others need to be possessed (cf. Vogeley, & Newen, 2002). The latter issue will be explored in the next part of the paper.

5. Differentiating perspectives

To try to understand this issue in greater depth, we need to reference developmental and clinical research on differentiating one's own and others' mental states. It seems that this ability is an important factor when it comes to adequate mentalization.

One of the first studies on this topic was based on the false-belief test. A researcher was playing with a child, told her a story and showed her two puppets. The story has the following schema: person A puts object X in place M. When person A leaves the room, person B takes object X from place M to place N. Then, the researcher asked the child: "When person A comes back, where will she look for object X?" (cf. Wimmer & Perner, 1983). Both this and subsequent research has shown that most four-year-old children understand that person B will look for object X in place M, as she did not see the moment of it being taken to place N. This type of understanding demands inhibition of pointing to place N and taking the perspective of person B.

Research on mentalization development between three and five years of age shows that understanding others' mental states is related to the ability to reflect on alternative representations (cf. Gopnik & Astington, 1988). Children begin to understand that different people may have different observations and knowledge on a given topic. Between the ages of three and five they realize more often that knowing something is based on experience and that beliefs may change. According to the authors of this study, acknowledging the differences between one's own mental states and those of others enables reasoning on one's own mental states (cf. Gopnik & Astington, 1988). Hence, this capability seems to be one of the basic competences enabling people to understand themselves and others properly.

The importance of this ability may be noticed in research on mentalization disorders, especially borderline personality disorder. People with this disorder experience problems with emotion regulation, are highly impulsive and self-aggressive, and have difficulties in interpersonal relationships (cf. Fonagy & Luyten, 2009). This disorder may be explored from different perspectives; only the mentalization aspect is considered in this paper.

People who suffer from this disorder have impaired ability to separate the perspectives of oneself and others (cf. Bender & Skodol, 2007; Blatt & Auerbach, 1988; Fuchs, 2007). Presumably, this is the effect of difficulties in reflective thinking and other cognitive impairments. As a result, such people react more impulsively, emotionally, and overinterpret the external aspects of communication (such as facial expressions or gestures) (cf. Fonagy & Luyten, 2009). An underestimated feeling of self-agency is also present. Suffering from these problems, people with borderline personality disorder have difficulty experiencing intentionality of feelings, thoughts and actions, and they also experience a loss of integrity, which consequently leads to inadequate mentalization (Fonagy & Luyten, 2009).

Impaired ability to differentiate the mental states of oneself and others may lead to two kinds of effects (cf. Fonagy & Luyten, 2009). Firstly, there is a high probability of projecting one's own feelings and thoughts onto another person. Person A may suppose that others feel anger at the same moment as he or she feels it. Furthermore, even if such a person would like to understand person B, take his or her perspective and simulate his or her mental state, person A would probably have a problem

with realizing that the experience of person B may be different from his or her concept of this experience. Thus, it can be stated that in the case of person A, there exists a difficulty in verifying one's own concept of another's experiences, which is the result of impaired ability to differentiate two perspectives.

Secondly, there is a risk of acquiring another's mental states. If there is a difficulty in separating the emotions of oneself and others, there is a high probability of becoming sad in the company of sad people or angry in the company of angry people. On the one hand, such a mirroring process is common for everyone, while on the other, reflectivity and the capability to differentiate emotions of oneself and others enables a person to control such mirroring, at least to some degree. A person suffering from borderline personality disorder has problems in this sphere and because of that, there is a risk of being overwhelmed by others' mental states (cf. Fonagy & Luyten, 2009). However, such problems are not present all the time. They are mostly present during high emotional arousal.

One might ask if in this case it is possible to ascribe mental states adequately. It is undoubtedly highly difficult – one's own mental states are often ascribed to others, and others' mental states – ascribed to oneself. In the case of a coincidence of occurrence of the same belief or emotion, there might be room for adequacy. However, it is hard to claim that this adequacy is the result of any cognitive process and not just coincidence.

6. Conclusion

The abilities to differentiate perspectives and separate one's own and others' mental states seem to be necessary conditions for adequate mentalization. The capability to simulate others' mental states – to put oneself in someone else's shoes – is important but at the same time not sufficient. Similarities between interacting people are a comparable situation – they may contribute to understanding one another, but they may also impede this process. The ability to differentiate the agency of mental states is an essential element for social cognition. It is also important to understand that particular mental states may be shared with others, but not all the time (cf. Fonagy & Luyten, 2009).

This ability enables adequate mentalization as a result of considering various perspectives and recognizing the possibility of the coexistence of various representations of reality. It is probably

not the only aspect of social cognition enabling adequate mental state ascription, but it seems to be an essential one. Further research could help us to better understand how we can try to answer questions such as: What is he thinking? What does she feel? What do they think about that?

References

Ames, D. R. (2007). Everyday solutions to the problem of other minds: Which tools are used when. In: B. F. Malle, S. D. Hodges (Eds.), *Other Minds: How Humans Bridge the Divide Between Self and Others* (pp. 158–173). New York: Guilford Press.

Barlassina, L., Gordon, R. (2017). Folk psychology as mental simulation. In: E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. Retrieved from https://plato.stanford.edu/entries/folkpsych-simulation/.

Bender, D. S., Skodol, A. E. (2007). Borderline personality as a self-other representational disturbance. *Journal of Personality Disorders*, *21*(5), 500–517.

Blatt, S. J., Auerbach, J. S. (1988). Differential cognitive disturbances in three types of borderline patients. *Journal of Personality Disorder*, *2*, 198–211.

Bogdan, R. J. (2005). Why self-ascriptions are difficult and develop late. In: B. F. Malle, S. D. Hodges (Eds.), *Other Minds: How Humans Bridge the Divide Between Self and Others* (pp. 190–206). New York: Guilford Press.

Boyer, P., Robbins, P., & Jack, A. I. (2005). Varieties of self-systems worth having. *Consciousness and Cognition, 14,* 647–660.

Decety, J., Grezes, J. (2006). The power of simulation: Imagining one's own and other's behavior. *Brain Research*, *1079*(1), 4–14.

Fonagy, P., Luyten, P. (2009). A developmental, mentalization-based approach to the understanding and treatment of borderline personality disorder. *Development and Psychopathology*, *21*(4), 1355–1381.

Frith, C. D., Frith, U. (2006). The neural basis of mentalizing. *Neuron*, *50*(4), 531–534.

Fuchs, T. (2007). Fragmented selves: Temporality and identity in borderline personality disorder. *Psychopathology, 40*, 379–387.

Gallagher, S. (2008). Direct perception in the intersubjective context. *Consciousness and Cognition*, *17*(2), 535–543.

Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading (Philosophy of Mind)*. Oxford: Oxford University Press.

Goldman, A. I. (2009). Mirroring, mindreading, and simulation. In: J. Pineda (Ed.), *Mirror Neuron Systems: The Role of Mirroring Processes In Social Cognition* (pp. 311–330). New York: Humana Press.

Goldman, A. I. (2012). Theory of mind. In: E. Margolis, R. Samuels, S. P. Stich (Eds.), *The Oxford Handbook of Philosophy of Cognitive Science* (pp. 402–424). Oxford: Oxford University Press.

Goldman, A., Mason, K. (2007). Simulation. In: P. Thagard (Ed.), *Philosophy of Psychology and Cognitive Science* (pp. 267–293). Amsterdam: Elsevier.

Gopnik, A., Astington, J. W. (1988). Children's understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. *Child Development*, *59*(1) 26–37.

Gordon, R. M. (1995). Simulation without introspection or inference from me to you. In: M. Davies, T. Stone (Eds.), *Mental Simulation* (pp. 53–67). Oxford: Blackwell.

Guerini, R., Marraffa, M. (2015). Mentalization and self-identity: Philosophical and psychological perspectives. *Rivista Internazionale di Filosofia e Psicologia, 6*, 196–197.

Hodges, S. D. (2005). Is how much you understand me in your head or mine. In: B. F. Malle, S. D. Hodges (Eds.), *Other Minds: How Humans Bridge the Divide Between Self and Others* (pp. 298–309). New York: Guilford Press.

Hyslop, A. (2016). Other minds. In: E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy*. URL = <https://plato.stanford.edu/archives/spr2016/entries/other-minds/>.

Krueger, J., Overgaard, S. (2012). Seeing subjectivity: Defending a perceptual account of other minds. *ProtoSociology: Consciousness and Subjectivity*, *47*, 239–262.

Mitchell, J. P., Macrae, C. N., & Banaji, M. R. (2006). Dissociable medial prefrontal contributions to judgments of similar and dissimilar others. *Neuron*, *50*(4), 655–663.

Newen, A. (2015). Understanding others: The person model theory. In: T. Metzinger, J. M. Windt (Eds.), Open *MIND*: 26(T). Frankfurt am Main: MIND Group.

Newen, A., Schlicht, T. (2009). Understanding other minds: A criticism of Goldman's Simulation Theory and an outline of the Person Model Theory. *Grazer Philosophische Studien*, *79*(1), 209–242.

Perner, J. (1991). *Understanding the Representational Mind*. Cambridge, MA: MIT Press.

Robbins, P. (2004). Knowing me, knowing you: Theory of mind and the machinery of introspection. *Journal of Consciousness Studies*, *11*, 129–43.

Russell, B. (1970). Analogy. In: Thomas O. Buford (Ed.), *Essays on Other Minds* (pp. 3–8). Urbana: University of Illinois Press.

Vogeley, K., Newen, A. (2002). Mirror neurons and the self construct. In: M. Stamenov, V. Gallese (Eds.), *Mirror Neurons and the Evolution of Brain and Language* (pp. 135–150). Amsterdam: Benjamins.

Vogeley, K., Newen, A. (2009). Consciousness of oneself and others in relation to mental disorders. In: S. J. Wood, N. B. Allen, & C. Pantelis (Eds.), *The Neuropsychology of Mental Illness* (pp. 408–413). New York: Cambridge University Press.

Waytz, A., Mitchell, J. P. (2008). Two mechanisms for simulating other minds: dissociations between mirroring and self-projection. *Current Directions in Psychological Science*, *20*(3), 197–200.

Wimmer, H., Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition, 13*(1), 103–12.

Joanna Barska
ORCID 0000-0003-3758-6116
University of Warsaw
»Artes Liberales« Academy

Music vs. Truth: Illustrative Music in the Context of Musical Aesthetics[1]

Abstract

The paper discusses the problem of musical rhetoric in the Renaissance and the Baroque. In the 16th century, the imitative qualities of music were developed through the concept of *imitazione della natura*. The relationship between word and music stabilized. Representation was a major function and was inspired by the ancient concept of merging words, harmony, and rhythm. In ancient times, music had become a key to metaphysics and an important educational tool. Thus, specific compositions were related directly to a specific state of the soul, characters and emotions.

The author presents the aforementioned concept in later times, i.e. in the Renaissance, when not only theorists but also composers, performers and listeners assigned musical-rhetorical figures to specific meanings. The paper extensively discusses examples of the musical application of different musical-rhetorical figures and how they are used. The author underlines, however, that despite the universalization of musical language, the rhetorical system in music is highly diverse and is subject to individual contextual interpretations.

Keywords: musical rhetoric, representation in music, aesthetics of music

Medieval symbols, Renaissance-Baroque rhetoric, 18th-century illustrative music, Romantic "psychologizing" – there have been many "dictionaries" in the history of the aesthetics of music that were meant to help a piece of music convey specific ideas. It is worth looking into at least a few of the tools for shaping musical narration, including musical rhetoric devices, leitmotifs, program music, and musical quotations.

The 16th century in music brought the development of an aesthetic of **imitating natural sounds** (*imitazione della natura*),

stabilizing the relationship between verbal and musical content. The representative function became an important aspect of a musical composition. Inspiration drawn from the ancient idea of appropriately combining words, harmony and rhythm (or rather: a desire to return to that idea) was supposed to increase the power of musical expression and thus also influence listeners' souls.

> Having emerged from it [the lost choreia – JB], for many centuries music became a metaphysical key – a field of learning that reveals the fundamental structures of being – as well as a major tool of teaching. Ethical theories documented from the 7th century BC, according to which music is governed by laws (*νόμοι*) that rigidly assign types of composition to specific states of the soul, characters and emotions, inspired thinking about music up to the modern age (Pseudo-Plutarch, 1992, pp. 24–25, Plato, 1960, pp. 132–134). Parallel to this, in Pythagorean circles there developed a metaphysical science of harmony and numbers whose nature – and thus the nature of the micro- and macro-cosmos – is revealed by music. It was this trend of thinking about music, developed by the Platonic school, that gave grounds for placing it in the *quadrivium* of the *septem artes liberales* system, next to geometry, algebra and astronomy (Madeyska-Pilchowa, 2003, pp. 113–114).

In the age of humanism, music was subordinated to text, and the purpose of a musical work was defined by the triad *docere – delectare – movere*, highlighting the ethical aspect of music (and its influence) (Zawistowski, 2016, pp. 2–3). This subordination manifested itself in imitation – treated as pure expression – of phenomena of nature (Zawistowski, 2016, p. 3). In a vocal-instrumental piece, word painting followed the principle of decorum: If the text described a sad, serious event, the composer was required to use appropriate musical means, such as slow rhythm and a melancholy tune, to express such a mood. Words important for the message were to be highlighted by means of a compatible interval pattern, melodic pattern, rhythm and key. Sound was supposed to express whatever the lyrics contained; this was how the duty of a composer was understood in the age of humanism (Lisiecki, 1993, p. 19, Wesołowski, 1979, p. 166).

This kind of practice can be observed, for example, in the later motets and masses of Josquin dés Prés. The *Credo* from *Missa Pange lingua* (1514/1515) includes musical-rhetorical figures typical of many masses and worth a closer look.

Example 1: first fragment from the *Credo* from *Missa Pange lingua* by Josquin dés Prés

Example 2: second fragment from the *Credo* from *Missa Pange lingua* by Josquin dés Prés

The first quote is the fragment speaking of God descending to earth. The words "[Qui propter nos homines] et propter nostram salutem **descendit de caelis**" ("who for us men and for our salvation **descended from heaven**") is accompanied by a descending movement in the melody. The next fragment is a reversal of this motion. The melody illustrating the words "Et ascendit in caelum" ("And ascended into heaven") is a rising one.

Another place worth noting is the fragment that comes after the caesura crowning the descension. The character of the music changes suddenly and perceptibly; vigorous rhythm and joyful harmonics give way to solemn phrases sung in a dignified, measured rhythm extended by fermatas. This happens on account of the following words: "Et incarnatus est de Spiritu Sancto ex Maria Virgine, et homo factus est" ("And by the Holy Spirit was incarnate of the Virgin Mary, and became man"). Once the mystical transformation of God into man has resounded, the piece's character and tempo return to the previous form.

This kind of symbolic articulation of moments of ascension, descension or transformation became a rule, returning in many

masses from later periods[2]. In significant fragments of madrigals, some composers illustrated key words according to a specific pattern indicating, for example, that the earth is linked to a low register, heaven to a high one, descending progressions denote dying or death, grace note runs with rich figuration mean fleeing, while the melody outline can express the movement of waves, or even the eyes of a beloved (two notes of equal length and pitch) (Zawistowski, 2016, p. 5).

The above-described ways of developing a work of music were followed in the 16th, 17th and even 18th centuries with the help of similar musical-rhetorical figures. These figures are a combination of music elements (melodic pattern, rhythm, tempo), the use of harmonic devices and pauses in a specific way, and also the intentional use of forbidden or incorrect – from the point of view of the composition rules of the time – harmonic solutions and melodic leaps. In some cases they serve as readymade music "models" that are meant to express very specific concepts, affects, or to imitate the intonation of speech. The composition process at that time was modeled on the arrangement of speech: Thanks to the choice of appropriate musical-rhetorical figures, music could become an equivalent for a given thought or emotion (Bartel, 1997).

Next to composers representing what was referred to as tone painting, who focused on details, on portraying certain phenomena and situations, the 16th century also included those who used figures expressing general moods, feelings or emotional states (Bartel, 1997, pp. 5–7). It also needs remembering that this kind of humanist approach to music coexisted with other trends and schools that put emphasis, for example, on decorative polyphony or coloristic qualities.

Late-Renaissance musical-rhetorical figures expressing natural phenomena and emotional states evolved in the Baroque period

---

[2] Zawistowski gives other examples of evident rhetoric being used in works by Josquin dés Prés, noteworthy ones including the motet *Absalon, fili mi* in which small interval distances throughout the piece as well as the intensive presence of small seconds and thirds, considered dissonance in the composer's time, express King David's pain and suffering following the loss of his son Absalom. On the other hand, in the well-known motet *De profundis clamavi*, which is based on Psalm 130, despite the well-developed polyphony, individual melodic phrases are clearly subordinated to the message whereas a slow tempo and dark tones evoke an atmosphere of prayer, longing, complaint, sadness, pleading. Larger intervals and consonant sounds appear with reference to hope for the absolution of sins (Zawistowski, 2016, p. 4).

into a complex rhetoric widely discussed at the time. The 17th-
and early 18th-century belief that the main purpose of music is
to arouse passions and affects yielded the **theory of affects** that
had an especially strong presence in the German-speaking world.
Stereotypical turns included a countless number of figures (*loci
topici*) that "depicted" or "illustrated" affects in music (Bukofzer,
1970, p. 524)[3]. Listing and describing musical figures, which
began in the 17th century, was one way of perceiving music in
a rhetorical context. The science of affects concerns a powerful,
brief emotional event whose occurrence is noticeable thanks to
physiological symptoms (red face, tears, accelerated heartbeat,
etc.). Baroque aesthetics – reflected both in theoretical thinking
and in composing and performance practice – was based on
expressing the system of affects by means of sound structures
and thus triggering those affects. "Baroque music," as Szymon
Paczkowski writes, "was meant to throw people off balance, pro-
voke them to extreme feelings, put them in a state of ecstasy"
(Paczkowski, 1998, p. 11). Theorists of the time described the
affective character of music intervals, scales and genres. A joint
formula of the "doctrine of affects" (of using typical musical
figures) was developed in Germany (*Affektenlehre*).

Every theorist proposed his own description of figures. In
the works of Jesuit scholar Athanasius Kircher (1601–1680) the
key, inspired by Pythagorean concepts, is a striving for simplicity
of intervals and their proportion. The closer this is to 1:1, the
closer it is to divinity and consonant sounds. Increasing the level
of complexity causes dissonant sounds, which in turn evoke pain
and suffering (Paczkowski, 1998, pp. 10–11). The functioning of
musical-rhetorical figures was described differently in 1739 by
Johann Mattheson, a German composer, singer and music theo-
rist[4]: Sadness should be expressed with a slow, heavy and sluggish

---

[3] It is worth remembering, however, that in the 17th and 18th centuries af-
fects were not an illustration of the composer's or performer's emotions, but an
attempt to depict the immanent traits of a given sound sequence; the painting of
violent emotions had all the hallmarks of stylization.

[4] "Since instrumental music is nothing other than a language of tones, or
eloquence in sound, it must always aim its intention at a certain movement of the
heart. To arouse this, it must take due care about the power of intervals, the deft
division of movements, the appropriate continuation, and other things of the
sort" (Dahlhaus, 2007, p. 28; English quote from Carl Dahlhaus, *Esthetics of
Music*, translated by William W. Austin, Cambridge University Press 1982, p. 24).

melody, hatred should be represented by unpleasant, rough harmony and an analogous melody. This theorist also characterized different dances, including *gigue* as an expression of ardor and desire, and *courante* as an expression of sweet hope and courage. These very imprecise explanations prove how difficult it was to unify the principles of the theory of affects. Nevertheless, despite the diversity of such attempts, it is hard not to notice they have many things in common. One example is chromatic second progressions, which were always linked to sadness, pain, suffering or grief[5]. Therefore, the dictionary of musical-rhetorical figures should not be treated as a uniform model that is not subject to interpretation, or as instructions (on how to read and write music). Because, every means of expression, obviously, should be interpreted in the context of how it functions in a specific composition, taking into account the composer's stylistic idiom. Let us consider an example. The same figure, i.e. vertical chords written out in a heavy, pointed rhythm, illustrating the flagellation of Christ, appear in the appropriate places in both Bach's *St. Matthew Passion* and Handel's *Messiah* (Zawistowski, 2016, p. 9). The former composer, however, approached individual words with greater precision, giving them an appropriate melodic outline and already at this micro-level taking advantage of the potential of harmonic sequences, which remain in a direct relationship with the lyrics of the piece. Handel, on the other hand, treated the text more freely and with greater flourish, proposing distinctive motifs and conventional melodic patterns that do not refer to the mood of the individual words (Zawistowski, 2016, p. 9).

There is no doubt that musicians of the late Baroque, especially in Germany, were not only aware of the theory of affects but very often also followed its principles in their compositions. In this way, the fundamental assumption of the science of affects, whose objective was to standardize music composition practice and universalize the language of music (the possibility of communicating through music), was largely fulfilled (Paczkowski, 1998, p. 14). In the 19th century this objectivization of the character of musical feelings gave way to expressing the self

---

[5] Another issue is the Baroque's linking of a given group of affects to specific keys as well as symbolic references to specific numbers; for example, the Holy Trinity can be invoked by a triple rhythm, a three-voice texture, a thrice-repeated distinct motif or a melisma, etc.

and an individual "language of feelings." Carl Philipp Emanuel Bach, Christian Friedrich Daniel Schubart and Samuel Gottlieb Heder, to name a few, rejected the theory of imitation as being trivial: A composer was not supposed to portray feelings but "to force out his selfhood in music" (Schubart)[6]. We should not forget, however, that Romanticism also developed its own symbolism, one that was no longer imitative in nature but evolved into a whole system of leitmotifs.

The system of musical-rhetorical figures developed in the course of three centuries is diverse and subject to individual contextual interpretation: analysis of form, texture and other elements of a work. At this point it is worth looking into at least a few examples of the most popular figures related to the aesthetics of affect imitation and showing how they functioned in specific compositions. A clear division of these is proposed by Chomiński and Wilkowska-Chomińska, who have distinguished two categories: **onomatopoeia figures** and **emphatic figures** (Chomiński, 1974, p. 92).

Besides their purpose of portraying certain elements of reality or highlighting the lyrics being sung, onomatopoeia figures have a major role in shaping the structural elements of a work of music. Compared to emphatic figures, which are more conventional in character (Chomiński, 1974, p. 103), onomatopoeia figures are easier to interpret when listening to music because their function is to liken (*assimilatio*) the musical structure to a selected sound phenomenon, therefore referring to a specific designatum (Chomiński, 1974, p. 103).

The fragments from the *Credo* from *Missa Pange lingua* by Josquin dés Prés quoted earlier, i.e. the popular *ascendus* (rising melodic outline) and *descendus* (falling movement of sounds), besides highlighting the movement of ascending and descending (in this case, in the sense of ascending to heaven, descending to earth), would have gained a wider, symbolic interpretation in later periods. Ascending could also refer to elevation, to lofty, wonderful and moral things, while descending could be tied to humiliation, meekness, obsequiousness, downfall, submissiveness, sometimes depression (Chomiński, 1974, Zawistowski, 2016, Paczkowski, 1998).

---

[6] It needs remembering, though, that we are speaking here of the "I" of musical expression, not a real person (the composer). See Dahlhaus, 2007, pp. 24–26.

The most popular (and earliest) onomatopoeia figures were those that imitated sounds made, for example, by animals (*assimilatio*): cuckoo, nightingale, rooster, dog. Apart from these, there were also sudden and violent sounds of battle, hunting, or a storm. The *tirata*, as this is called, is a series of sounds running up or down that can suggest a thunderbolt, throw, projectile, attack, etc. In Bach's *Passions* it occurs in the scene of the temple veil being torn in two and the scene of the earthquake. And here, again, we need to underline the role of the verbal context, which in some cases is essential for identifying the function of a given run or figure.

In Orlando di Lasso's *In hora ultima*, the sound runs illustrate individual words: "In hora ultima peribunt omnia: ... jocus, risus, saltus, cantus et discantus" ("At the last hour all things shall perish: ... jesting, laughter, dancing, song and descant"). One interesting fragment is the motif of laughter, which can only be identified thanks to the word "rirus" on which melismas imitating laughter are performed.

Example 3: Orlando di Lasso, *In hora ultima*

It is similar with the figure of the fugue in early music, where the first voice is imitated by successive ones, thus expressing pursuit, flight; another example is the figure of the cross, which is only legible upon analysis of the notation, i.e. a series of four notes that form a cross on the staff (e.g. e-b-B-e)[7]. In subsequent

---

[7] Musical symbolism is a much broader issue, and is subject to interpretation. Thanks to this, Albert Schweitzer was able to link the structure of Bach's Prelude and Fugue in E-flat Major BWV 552 to the Holy Trinity, justifying his idea with the fact that the key signature of both pieces has three flats, the prelude comprises three sections, and the fugue is a triple one.

centuries, as a result of increased interest in the musical form itself and in emotional expression, onomatopoeia devices lost their importance, although they did sporadically appear in various compositions[8].

Wherever composers wanted to achieve an accumulation of the **emotional aspect**, initially in madrigals and then in operas, cantatas, masses and dramatic laments[9], we find a gradual succession of halftones going up or down, which was the main way of depicting suffering. Pathopoeia, as this is called, is among the most frequent sound structures appearing in the context of **emphatic figures**. The interpretation of these figures is more ambiguous than that of onomatopoeia ones, and depends more strongly on the context. To a considerable extent, these figures take advantage of the evocative potential of melic structure and are developed by building appropriate intervals – e.g. diminished and augmented intervals, second progressions, or, on the contrary, greater intervals between sounds. Agogic or evocative contrasts, intensification or embellishment – these are the sort of means that could be used to underline the importance of a given word or phrase.

Means that use a different type of expressiveness include silence (*aposiopesis*) and the sigh (*suspiratio*). The significance of these devices is verified by the context of the dramatic situation.

---

[8] Cf. Chomiński, 1974, p. 104. In Joseph Haydn's Symphony No. 18 in G Major "Le soir" the finale is a storm ("La tempesta"), which also appears in other works by the same composer, illustrated with abrupt contrasts at the level of dynamics, agogics and instrumentation (juxtaposition of the *tutti* fragment, i.e. played by all the instruments, and the solo fragments played by the violin, flute, cello). Elements of painting with music can also be found in Beethoven's symphonies. Symphony No. 6 in F Major "Pastoral" includes a movement called *Scene by the Brook* in which sounds representing different birds are marked: nightingale, quail and cuckoo. The storm motif that was popular in the 18th century appears in the scene called *Thunder, Storm* with short motifs accented by the orchestra (increased to include trombones), harmony parts, a string tremolo, figuration by cellos and double basses, an intensive dynamic and chromatics augmenting the building tension. Different tools were used by 19th-century composers, who chose other areas in which to use the expressiveness of orchestral language and associations evoked by musical structures. More examples could be offered, leading us all the way to the 20th and 21st centuries. For more on this, see Chomiński, 1974, pp. 453–460.

[9] See works by Claudio Monteverdi: selected madrigals or e.g. "Lasciate mi morire" from the opera *Arianna*. Here, the composer uses the evocative qualities of melic structure, the special expressive values of small sound intervals.

For example, in Gluck's *Iphigénie en Aulide*, in the farewell scene between Iphigenia and her mother, Clytemnestra's voice gradually breaks down and ultimately dissipates (for more on this, see Chomiński, 1974, p. 95). *Aposiopesis* also appears as a figure of death, for example in passions in the moment when Christ gives up his spirit. *Suspiratio*, meanwhile, is used in an interesting way by Monteverdi. "Farewell, Rome, my fatherland, my friends" is how Nero's wife Octavia says goodbye to Rome in *L'Incoronazione di Poppea* (1643). Having fallen in love with another woman, the emperor is sending her into exile. The word "adio" ("farewell") is broken up with the help of pauses, strengthening the dramatic effect. The inability to utter the word "adio" creates an impression of deep despair, pain, powerlessness. The *suspiratio* is duplicated by the basso continuo, which – using sparing means – underlines the words through emphasis, additionally intensifying the expressiveness of this musical statement.

Example 4: fragment from the aria "Adio, Roma!" from Claudio Monteverdi's opera *L'Incoronazione di Poppea*

Emphatic devices include all those whose purpose is to distinguish (by means of dynamics, agogics or embellishment) a given fragment of the lyrics (*accentus*), for example by using an intentional technical mistake or false note, or by repeating a given word. The *anaphora* is a figure in which the motif is obsessively repeated in the bass (as is the case in Baroque laments, passacaglias and chaconnes), or in which a brief motif is taken up by individual voices of the polyphonic texture. This is what happens in the chorus in Bach's *St. Matthew Passion* in which the disciples, over and over again, anxiously ask Jesus, who has announced one of them will betray him, "Is it me?" ("Bin ich's?"). The *anaphora* becomes an expression of fear and curiosity. At the same time, this motif is a rhetorical question (*interrogatio*), a measure imitating speech, with rising intonation.

Example 5: fragment from Johann Sebastian Bach's *St. Matthew Passion*

For composers, performers and discerning listeners of the 16th, 17th and early 18th century, musical-rhetorical figures were the foundation of a clear, conventional musical language (Chomiński, 1974, p. 109). Apart from the most popular and most strongly conventionalized means, musical-rhetorical figures do not form a dictionary of topoi with stable meanings and functions, and therefore should be interpreted in the context of music and lyrics. Importantly, familiarity with the rhetoric used in a given work enables performers to execute it in accordance with the musical notation, choose the appropriate articulation, etc. Nikolaus Harnoncourt compares the musical language of a given period to a foreign language that needs to be learned, with its own vocabulary, grammar, articulation (Lisiecki, 1993, p. 17). Something that was once obvious to audiences, today – according to many researchers – should be reconstructed in such a way that in analyzing historical structures, musicologists should refer to the historical culture and not just make use of contemporary descriptive tools (Lisiecki, 1993, p. 14). When the aesthetics of affects lost its popularity and reach, the semantic contexts of musical-rhetorical figures lost their clarity as well. Some of them, however, were and are transformed, appearing in musical works to this day and fulfilling different expressive functions. That is why I believe reading a musical text in terms of the rhetorical devices used within it is useful as an analytical tool, and in some cases (especially music from the 16th to the mid–18th century) is even an essential context for a given

communication situation. However, I do not consider this to be a closed system of readymade "keys to interpretation" but allow room for a diverse range of interpretations.

References

Bartel, D. (1997). *Musica Poetica: Musical-Rhetorical Figures in German Baroque Music*. London: Lincoln.

Bukofzer, M. (1970). *Muzyka w epoce baroku. Od Monteverdiego do Bacha* (Polish edition of *Music in the Baroque Era. From Monteverdi to Bach*, transl. E. Dziębowska). Kraków: Państwowe Wydawnictwo Naukowe.

Chomiński, J., Wilkowska-Chomińska, K. (1974). *Formy muzyczne* [Music Forms], vol. 3: *Pieśń* [The Song]. Kraków: PWM.

Dahlhaus, C. (2007). *Estetyka muzyki* (Polish edition of *Musikästhetik*, transl. Z. Skowron). Warszawa: Wydawnictwa Uniwersytetu Warszawskiego.

Lisiecki, W. (1993). Vademecum muzycznej „ars oratoria" [Handbook of the musical "ars oratoria"]. *Canor*, *3*(6).

Madeyska-Pilchowa, A. (2003). Retoryczność psalmów Mikołaja Gomółki [The rhetorical nature of Mikołaj Gomółka's psalms]. In: M. Hanczakowski, J. Niedźwiedź (Eds.), *Retoryka a tekst literacki* [Rhetoric versus Literary Text] (pp. 112–134), vol. 2. Kraków: Universitas.

Paczkowski, Sz. (1998). *Nauka o afektach w myśli muzycznej I połowy XVII wieku* [The Science of Affects in Musical Thought of the First Half of the 17th Century]. Lublin: Polihymnia.

Plato (1960). *Prawa III* (Polish edition of Plato's *Laws III*, transl. M. Maykowska), 700B–701C. Warszawa: Państwowe Wydawnictwo Naukowe.

Pseudo-Plutarch (1992). *O muzyce* (Polish edition of *De Musica*, transl. K. Bartol). Wrocław: Ossolineum.

Wesołowski, F. (1979). Wprowadzenie do retoryki muzycznej [Introduction to musical rhetoric]. *Zeszyt Naukowy Akademii Muzycznej we Wrocławiu*, *21*.

Zawistowski, P. (2016). *Rozważania na temat retoryki w muzyce baroku* [Reflections on rhetoric in music of the Baroque]. Retrieved from http://chopin.man.bialystok.pl/umfc/wp-content/uploads/2016/04/02-02.pdf.

Karolina Kolinek-Siechowicz
ORCID 0000-0003-3939-5276
Faculty of "Artes Liberales", University of Warsaw
»Artes Liberales« Academy

Truth and Early Music: The Intersection of Arts and Humanities

Abstract

The aim of the paper is to show how the concept of truth relates to the Early Music revival. The author shows consecutive stages of the revival in the light of different problems related to the tension between the arts and academic knowledge. The development of the movement is compared with the general concept of music revival presented by Bithell and Hill (2014).

Keywords: Early Music, revival, Historically Informed Performance, authenticity, truth

The aim of this paper is to describe how the concept of truth relates to the Early Music movement and what it says about the relation between truth and both the arts and the humanities.

Before explaining this problem, I would like to make some important terminological distinctions. First of all, I will make a distinction between truth *in* music and truth *about* music. Then I will define what I mean by the term "Early Music." After that, I will be ready to consider the problem of truth *about* music in the light of the Early Music movement which is an especially interesting case of applying the truth concept in music studies.

1. Truth in music

The problem of truth in the case of music is usually considered in the light of musical content or meaning. Although music is a nonrepresentational art and does not correspond to the actual world in the manner language does, the issue of truth and music

seems to be quite well developed both by musicologists and philosophers of music. The tradition of ascribing extra-musical meaning to the sounds of music may be traced back as early as Plato, who codified the influence of particular musical modes on human morality. Every musical epoch had its own system of musical meaning which, in some cases, is also transparent to us thanks to knowledge derived from theoretical sources of the past. However, the tendency to assign verbal narrative to a piece of music flourished in the 19th century when instrumental music began to dominate the scene. The vanishing of words from a musical composition's structure opened the door for less restrained interpretation and invited philosophers to treat music as the highest sublimation of the spirit. Today, as the successors of the Romantic mode of listening, we still meet with such an approach. It is clear in music criticism, but it is also valid for the contemporary philosophy of music, which – while seeking an explanation for music meaning – takes as examples almost exclusively 19th-century repertoire, with some exceptions from the 18th and 20th centuries (Goehr, 1992). Such a discourse belongs to a way of thinking about music that is rather ahistorical, emphasizing its universal and timeless values and not caring about historical accuracy.

2. Truth about music

While philosophers explore truth *in* music, music historians and historically aware performers tend to speak of truth *about* music instead. At first glance, this question seems to be closer to the problem of historical truth, which is an inherent matter of concern for history as an academic discipline. It is even more evident when we realize that both the history of music and Historically Informed Performance rely mainly on historically proven knowledge about the musical past. However, the problem of truth considered in the light of music history and performance is situated at the point of intersection of two issues: the issue of historical truth and the issue of truth in (or about) the aesthetic experience. That is what makes the issue of truth about Early Music so complicated – it is never limited to the application of any theory about truth only to the theoretical (historical) or just to the performative (aesthetic) aspect of this phenomenon.

That is also what makes this issue so interesting – it allows us to confront the problem of our attitude to historical facts and their contemporary interpretations while being thrown into experiencing art which, by its nature, eludes an approach of simply judging facts. Therefore, analyzing the problem of truth in the case of the Early Music movement may be helpful for seeing mutual connections between the arts (i.e. music) and the humanities and ways of reconciling them in the field of performance.

3. Early Music borders

The issue of Early Music's historical borders is not obvious, as regards both defining its "earliness" and pointing out when exactly the idea of playing non-contemporary music emerged. To make the problem a little clearer, John Haines in his article "Antiquarian Nostalgia and the Institutionalization of Early Music" (2013) distinguishes early music as the general concept of music created before the present time (which may apply equally to the reality of the 16th and the 21st century) from Early Music as a modern concept, conceived in the 20th century and strictly connected with musical institutions which were born during this time. It is important to note that many attempts to bring repertoires from the past back to life were made not only in recent centuries, but also much earlier. That is the casus which Haines relates to general Renaissance antiquarianism.

In this paper, while considering the problem of truth about music, I will concentrate on the latter meaning of the term proposed by Haines – the Early Music movement, also known as the Historically Informed Performance (HIP) movement, which appeared at the beginning of the 20th century and was based on academic research on music history. The aim of HIP was to reconstruct music of the medieval, Renaissance and Baroque eras (later including Classicism and Romanticism as well) in the way it originally sounded, by analyzing primary sources, bringing back to life old instruments, and discussing problems such as composers' intentions, historical styles and the extra-musical context of particular repertoire. As a social and aesthetic phenomenon, the emergence of HIP is also often regarded in the light of the shift from a Romantic to a modern approach to music making and listening. A concise definition of Early Music that concentrated

on social significance was provided by Laurence Dreyfus in one of the earliest meta-reflective articles concerning this phenomenon (1983, p. 298):

> It is therefore more useful to define Early Music as a late twentieth-century ensemble of social practices instead of restricting it to the works which occasion the interest. To be blunt, Early Music signifies first of all people and only secondarily things.

Restricting the subject of this paper to HIP may help to look closer into the problem of the relationship between truth, the arts and the humanities in the case of music. The Early Music movement seems to be a unique musical phenomenon which engages a great deal of effort in historical research while not evading the live performance of the examined repertoire even when full evidence of its every dimension is unavailable.

Describing selected stages of the Early Music movement in the 20th century, I would like to demonstrate how the balance between striving for authenticity and understanding modern music aesthetics influenced the issue of truth about music.

4. True tension of revivals

It is quite difficult to point to the exact moment when the need to get to know the whole truth about music from the past became irresistible. We can observe several signs of a disposition to revive the musical past in different periods of history. It is often claimed that the first large-scale music revival was the Florentine Camerata's activity at the end of the 16th century. The idea of reconstructing ancient drama, with its music which served the word, was the main driving force for the creative output of composers and poets centered around Count Giovanni de' Bardi. Haines (2013) links this movement to the general antiquarianism trend which was the fuel for Renaissance interest in the ancient past. It is important to understand that both the lack of evidence for reconstructing music and the impact of contemporaneous aesthetic assumptions resulted in a process of recreating (with the emphasis on *creating*) rather than reconstructing – the end result of those experiments was the birth of a new genre which today is known as opera.

The reason why the Florentine Camerata is often mentioned in the context of the Early Music revival is the observation that

each attempt to reconstruct music of the past ultimately results in the development of new qualities of contemporary performance, even though the movement's creators believe they *really* grasp what is now gone. This incompatibility of imagined ideas and effects is universal for all music revivals (including not only early music but also folk music revivals). In fact, the process of reviving music requires respect for both the truth about mute music sources of the past and the sound of music performance of the present. The tension between what musicians *know* and what they *create* illustrates the difficulty of judging a performance which aspires to historical appropriateness. However, the ratio between faithfulness to musical sources and artistic creativity changed throughout successive decades of the 20th century. Taking those changes into consideration may help to distinguish what was (and what is) regarded as truth about music and what were its conditions.

5. True instrument

The symbolic wellspring of the modern Early Music movement is often regarded in the light of the activity of such musicians as Wanda Landowska and Arnold Dolmetsch. Their interest in historical instruments and efforts aimed at their reconstruction mark them out as pioneers of an approach that respects not only written sources, but also the use of particular sources of sound which had disappeared by the 19th century. Although each of them represented a different attitude toward the problem of reconstruction, they both shared the belief that modern listeners should get to know forgotten musical repertoires and the sound of historical instruments.

Of course one should not forget that Landowska's harpsichord made by Pleyel was a kind of hybrid instrument that combined the mechanism of the harpsichord and the piano. One should also keep in mind that she admitted herself that she was far from the composers' intentions, as she had to interpret their music within her own sensibility (Landowska, 1964). But, at the same time, she was convinced that her contemporaries should rediscover the repertoire of the Middle Ages, Renaissance and Baroque, and acknowledge its importance in culture as much as they praise the paintings of the Old Masters. Landowska was

also a strong opponent of the idea of progress in music, which she regarded as harmful for the modern reception of music of the past (Landowska, 1964).

Although she understood that knowledge of sources is crucial for interpreting early music, she also claimed that the music should attract listeners. That is why she believed that her massive Pleyel instrument was more appropriate for modern concert halls than the original harpsichord, which is why she claimed that through her sensibility and emotional involvement in Bach's works (among others) she can attain their proper character. Landowska's strong artistic personality not only helped to rediscover the harpsichord as the medium for the music of Rameau, Scarlatti or Bach, but also caused modern composers to be attracted to this instrument. Thus, the harpsichord became a real part of early 20th-century culture, benefiting both from the Early Music revival and modern music composed for it by composers drawn to its exotic sound.

There is no doubt that in the light of later developments, Landowska was far from authenticity in her performances. However, she was extremely successful in introducing the harpsichord into modern concert halls and drawing people's attention to music that had not been heard for centuries. Her numerous writings about early music show that she kept abreast of new musicological discoveries. She also studied 18th-century treatises about the art of playing the harpsichord, but still her approach to historical appropriateness was quite liberal.

6. True song

At the beginning of the 20th century, German musicologists such as Friedrich Ludwig conducted fruitful research into medieval polyphony. Their discoveries did not wait long to be performed. Anna Maria Busse Berger (2019) describes the process of awakening an interest in the Middle Ages in the second and third decades of the 20th century, which had its impact also on musical preferences. German musicological scholarship at the time was very close to initiatives that included performing medieval repertoire. The most spectacular such movements were Wandervogel, Sing- and Musikbewebung. Their attitude had the hallmark of a utopian vision of a time unspoiled by technological progress and appealing in its simplicity and closeness

to nature. This imagined innocence of the past was something that attracted people in the same way as non-European cultures did. However, it was exotic enough to serve the concept of "the past as a foreign country" (Lowenthal, 1985) and, at the same time, familiar enough to be identified as the root of contemporary Western (i.e. German) art tradition. What it is particularly interesting is the fact that those brotherhoods promoted active music making, exactly in an age when the mechanical reproduction of music began to develop and progressed rapidly. Returning to quasi-amateur performance was undoubtedly a way to experience music more directly, perhaps even more genuinely. Busse Berger also mentions the great musicologist, Heinrich Besseler, who was one of the first to underline the chasm between the premodern and early modern era, visible also in the shift from active music making to passive contemplation.

The case of the close relationship between German youth movements and the beginnings of musicology as an academic discipline also shows the importance of scholarship in seeking "historical truth" in music performance. It is not a coincidence that early musicological research, looking for new territories to investigate, looked into early music. Undoubtedly this was a promising area of research – unexplored, puzzling and, at least at first, regarded as almost purely theoretical. However, as the later history of the Early Music movement shows, discoveries of that time were essential for performers who finally made this music live again.

7. True notation

Collaboration between scholars and performers seemed to be indispensable for fulfilling the aims of the Early Music revival. Although many musicians also had an academic (historical and analytical) background, musicological scholarship itself helped to methodize and clarify many issues which were crucial for performance. However, this relationship was complicated rather than happy and trouble free. Initially, both scholars and performers tended to disbelieve each other's discoveries. There are plenty of interesting opinions which prove that the Early Music movement was perceived not only as a refreshing novelty, but sometimes also as a dangerous misinterpretation. The most widespread critiques

of the Early Music movement were those of T.W. Adorno, who saw hope for the future only in the avant-garde (1967), and then of Richard Taruskin, who believed that HIP aesthetics is mainly a symptom of modern taste (1995). But there were more opponents, to mention Frederick Neumann who claimed that the Early Music movement was a "childhood disease … caused by a somewhat naive trust in the infallibility of historical treatises, the symptoms of [which] are manifested in a faulty interpretation of these documents" (Dreyfus, 1983).

Although Neumann's words are bitter and certainly not impartial, they are also significant. The problem of misinterpretation of sources was indeed the Early Music movement's real cul-de-sac. The longing for "authentic" performance, imagining the composer's intentions and believing the letter of the notation, led to what was undoubtedly too narrow and misleading a vision of the musical notation. The most strict stage of the HIP movement was often called "music philology," and assumed that what a composer wrote down is a kind of sacred testimony which should be protected from the anachronistic visions of editors. However, composers' manuscripts, first editions and sketches of particular pieces turned out to be insufficient to grasp the performance practice of the past, as many issues were simply not included in the scores. Without too much simplification, we can say that the older the piece was, the more features were unwritten, as oral tradition was the main way of teaching how to make music.

At the same time, approximately until the last decades of the 18th century the only performed music was contemporary music, so scores were rarely intended to serve future generations. The idea of performing music of the past in the way it may have been heard by its first listeners appeared much later – too late to re-establish the oral part of performance practice.

That is why "straight" performances (Taruskin, 1995) or, in other words, performances which set aside the performers' expression and try to "read" the score as it is, were so often criticized. One valuable attempt to re-evaluate historical and modern performance was an article by Daniel Leech-Wilkinson (1984), who compared "authentic" and "inauthentic" performances, underlining that sometimes it may be better to remain not historically informed than play music mechanically, without any signs of personal engagement.

More recent HIP attempts seem to be more balanced in their trust in the notation. On the one hand, contemporary scholars and performers have incomparably wider access to original sources, which are digitized and often placed in the public domain (e.g. The Josquin Research Project at Stanford University); on the other hand, they are becoming more aware of the importance of oral tradition and many extra-musical sources that also can say a lot about how the music might have sounded.

8. True reception

Although the Early Music movement puts performance practice at the focus of interest, this is not its only concern. There is no doubt that any research considering historical issues cannot avoid the problem of choosing the perspective of description – no historical narration is absolutely objective. The issue of early music's reception is complex, as almost every epoch had its own vision of the history of music. One of the most symptomatic examples is the reception of J.S. Bach's music. The 19th-century Bach revival is often regarded as the symbolic onset of the Early Music revival. In many publications that describe the latter, the Berlin Singakademie's performance of Bach's *St Matthew Passion* conducted by Felix Mendelssohn in 1829 is marked as the cornerstone of reintroducing music of the past to the broader audience. Admittedly, Mendelssohn's performance is remarkable for many reasons, but its later interpretation became misleading for subsequent generations. Ellen Exner in her paper delivered at The UMass Amherst Bach Festival and Symposium (2019) presented a condensed story about Bach's reception before 1829 and the mythological potential of Mendelssohn's revival which was developed by Romantics into a historically misleading but culturally appealing and useful story. Such a myth was also received by Adorno and interpreted through his own vision of historiography, resulting in the famous tendentious essay *Bach Defended against His Devotees* (1967).

Another example is the contemporary perception of J.S. Bach as a visionary composer who created his music for unknown generations. Such a view is shared by many music lovers and performers, who treat Bach's music as a universal domain of pure beauty and transcendence. However, many musicological

studies prove that Bach was a man of his time, totally immersed in the contemporaneous body of knowledge. Michael Marissen in his book *Bach & God* (2016) shows how inextricably Bach's music (even the secular) was bound up with current theological considerations and how deeply the composer's art was involved in them.

These examples only show a tiny section of the Early Music revival and its narratological implications, but are certainly quite spectacular, as Bach's music is a classic example of the tug of war between the mainstream and HIP within the classical music world. The Early Music movement's other shifts of narration seem less exciting but are no less important. A good example is the interest in the music of minor masters, underestimated by mainstream culture and brought back to listeners' attention by HIP. Discoveries of a similar type, which sometimes turn out to be uncomfortable for scholars and performers, come from re-attribution of famous pieces. These kinds of narration shifts always lead to further questions about the real understanding of authenticity and historical truth in the evaluation of musical aesthetics.

9. True revival

The newest publications concerning the Early Music movement often regard it in the light of theories of revival, situating the phenomenon in a wider context. While the first reflections about HIP were often limited to the problem of tensions between "new" and "old" performance paradigms within the classical music world, nowadays they are also built into a broader consideration of general tendencies toward revivals in modern society. One significant example of such an approach is the volume *The Oxford Handbook of Music Revival* (Bithell, Hill, 2014), which considers mainly ethnomusicological issues but also gives attention to the Early Music movement (Haines, 2014). In the first article in the collection, Bithell and Hill (2014) offer a list of six features common for any kind of musical revival. They can also be successfully applied to the Early Music revival, and such application may be refreshing for the overall reflection on the cultural and social circumstances of the movement. Enumerating them will also be a good summary of the issues outlined in earlier sections of this paper.

First of all, Bithell and Hill point to dissatisfaction with the present and a desire to make a cultural change. In the case of the Early Music revival, this kind of dissatisfaction is often interpreted in the light of the exhaustion of Romantic aesthetics and looking for new (old) means of expression. This is exactly where Taruskin (1995) looks for the reasons why the Early Music movement was born, when he describes it as the result of the 20th century's shift of musical taste. Also Butt (2004) and Haynes (2007) acknowledge the implication of HIP in surfeit of the stylistic means of Romanticism, especially when they juxtapose performance practice in both paradigms.

The second feature of musical revival is "identifying musical elements and practices as old, historical, or traditional, and determining their value [which] often involves selecting from or reinterpreting history and establishing new or revised historical narratives" (Bithell, Hall, 2014, p. 4), and that is what is described in section 8 of this paper.

The third facet is recontextualization which is inevitable when it comes to performing music of the past in present conditions. Recontextualization of early music fulfills all the types pointed out by Bithell and Hall, that is: temporal, geographical, and/ or social. Still, there is another type to add: an aesthetic shift which is closely related to the social shift but also needs to be considered. Musical repertoire of the premodern era was part of a world that did not know a purely aesthetic attitude, as art was a part of social practices and was perceived in this context. As many writers suggest (Goehr, 1994; Kivy, 1995; Berger, 2000), this is one of the most important effects of the modern divide which brings the institutions of the museum and concert hall as well as the idea of art for its own sake, resulting in an attitude of silent contemplation also in the case of music. After such a change of perceptual habits there is no return to performing and listening to music as if it were inseparable from its original context (a good example is performing Bach's *St. Matthew Passion*, originally intended to be part of the Good Friday liturgy, in a giant concert hall, with an intermission after the first part of the performance). But there have been several attempts not only to reconstruct the sound of early instruments, but also to put early music in the new light of contemporary social circumstances to restore its extra-musical significance. Such cross-temporal interpretations are common in the field of opera, where

the staging is often up-to-date while the performance may be historically informed.

Next Bithell and Hall point to the problem of the legitimacy of culture-bearers. In the case of early music, this seems to be a complex issue, as there is no direct path of inheritance for music that is sometimes more than 500 years old. However, many ethnomusicological studies show that there was a need to prove that some traditions were preserved in oral cultures and are now able to serve as a source for HIP performers who can learn from them, for example, how to sing a medieval chant (e.g. the famous CD entitled *Chant Corse* by Marcel Pérès and Ensemble Organum). Further examination showed that this was often the effect of the Eurocentric illusion that the more "primitive" a culture seems, the more untransformed is its musical tradition. Another way to legitimacy for Early Music performers is their scholarly background. Even if a performer is not educated as a music historian, they often need to show how much they know, writing books that combine historical knowledge and performance experience (Sir John Eliot Gardiner may be a perfect example, but Landowska can also be mentioned here). Where there is no obvious way to examine the historical correctness of a performance, this is how performers earn listeners' confidence in the authenticity (not only beauty or virtuosity) of their interpretation.

The fifth aspect is the revival's infrastructure, which includes "festivals, competitions, educational institutions, organizations, government policies, recording and distribution companies, and so on" (Bithell and Hill, p. 4). The Early Music revival, conceived as a rather informal and out-of-institution movement (Kelly, 2011), is currently an institutionalized part of the music industry and education, and its importance is still growing. Historically Informed Performance is no longer a niche – this becomes clear especially when we analyze its strong presence at "mainstream" festivals and in the recording industry, as well as the bumper crop of period instrument competitions (a good example is the recent establishment of the International Chopin Competition on Period Instruments).

All of these transformations lead to the last phase of the phenomenon: "post-revival," as Bithell and Hill call it. The Early Music revival undoubtedly reached this phase at least at the beginning of the 21st century, and we can still observe its effects in almost every part of the classical music world.

10. True conclusion, although the subject is truly inexhaustible

The dynamism of the Early Music revival in the 20th century not only fits the general theory of music revival, but also shows the tension between historical truth and artistic creation, or – to put it differently – between knowledge and the imagination. Several of the problems discussed above confirm that the Historically Informed Performance movement raised many important questions concerning authenticity, originality, fidelity to composers' intentions and, finally, the boundaries of performers' freedom, which are valid not only for the classical music world but also for the humanities, especially those disciplines which face the problem of historical truth. There is no doubt that the evaluation criteria for musical performances and academic research are not the same. However, awareness of the possible applications of what we know from historical research allows us to breathe life into the cultural practices of the past and include them in the contemporary art world. Understanding the role of the humanities not only as guardians of the heritage of the past, but also as guides in its implementations in the present, may help us recognize their value yet again and save their importance in academic structures and common appreciation.

References

Adorno, T.W. (1967). Bach defended against his devotees. In: *Prisms*, transl. S. Weber, S. Weber Nicholsen (pp. 135–146). London: Neville Spearman.

Bithell, C., Hill, J. (Eds.). (2014). *The Oxford Handbook of Music Revival*. New York: Oxford University Press.

Butt, J. (2002). *Playing with History: The Historical Approach to Musical Performance*. Cambridge: Cambridge University Press.

Berger, K. (2000). *A Theory of Art*. New York: Oxford University Press.

Busse Berger, A.M. (2019?). *Jugendmusik- und Singbewegung* – a chapter of a forthcoming book, courteously shared with me by the Author.

Dreyfus, L. (1983). Early music defended against its devotees: A theory of historical performance in the twentieth century. *The Musical Quarterly*, 69/3, 297–322.

Exner, E. (2019). Friedrich Smend, Friedrich Blume, and the Myth of Mendelssohn's "Matthew". A paper delivered on 04/13/2019 at The UMass Amherst Bach Festival and Symposium, Amherst, MA, USA.

Fabian, D. (2001). The meaning of authenticity and the Early Music Movement: A historical review. *International Review of the Aesthetics and Sociology of Music*, *32/2*, 153–157.

Goehr, L. (1992). *The Imaginary Museum of Musical Works: An Essay in the Philosophy of Music*. Oxford: Clarendon Press.

Haines, J. (2013), Antiquarian nostalgia and the institutionalization of early music. In: *The Oxford Handbook of Music Revival*. Retrieved from https://www.oxfordhandbooks.com/view/10.1093/oxfordhb/9780199765034.001.0001/oxfordhb–9780199765034-e-011. New York: Oxford University Press.

Haynes, B. (2007). *The End of Early Music: A Period Performer's History of Music for the Twenty-First Century*. New York: Oxford University Press.

Kelly, T. F. (2011). *Early Music: A Very Short Introduction*. New York: Oxford University Press.

Kivy, P. (1998). *Authenticities. Philosophical Reflections on Musical Performance*. Ithaca: Cornell University Press.

Landowska, W. (1964). *Landowska on Music,* D. Restout (Ed.). New York: Stein and Day.

Leech-Wilkinson, D. (1984). What we are doing with early music is genuinely authentic to such a small degree that the word loses most of its intended meaning. *Early Music*, *12/1*, 13–16.

Lowenthal, D. (1985). *The Past Is a Foreign Country*. Cambridge: Cambridge University Press.

Taruskin, R. (1995). *Text and Act: Essays on Music and Performance*. New York: Oxford University Press.

Anna Chęćka
ORCID 0000-0002-9340-7672
University of Gdańsk

Truth Embodied
in Music

Abstract

This contribution to the discussion on truth in music shows two ways of
thinking most often chosen by representatives of the contemporary Eng-
lish-speaking world's analytic philosophy. The first way assumes seeking truth
or falsehood in the actual musical work and its structures, while the other is
related to truth in experiencing music, and thus to the sincerity, authenticity
of a performance (and especially performers' faithfulness to the score and to
themselves). The text concludes with a literary example of the debate on truth
in music. For aestheticists, it can be a source of reflection on metacritical
issues of music performance, and for critics – encouragement to verify their
own interpretative strategies.

Keywords: music, truth, interpretation, value

When a philosopher considers truth in music, he or she
usually chooses two ways. One takes the form of speculation
about whether sound structures can correspond to true or false
statements. The other way suggests searching for truth in the
experiencing of a musical work that the listener attends thanks
to a performer. In the latter case, it often comes down to the
question of how to reconcile the performer's individual and
unique expression with attention to accurate reading of the
composer's text. This is related, for example, to subtle tensions
between fidelity of musical translation and betrayal, through
which performers may intuitively accentuate their own artistic
personality. It should be noted that in the case of classical music
performance, the reverse situation (doubting whether world-class
artists have anything authentic, honest, true to present through
their playing or singing) is definitely less common.

In my brief considerations, I would like to distinguish two examples of following these paths, on the basis of the English-speaking world's analytic philosophy. The third example, for a change, will not be scientific but literary. It will be shown, however, that behind a fictional character described in a book hides a real artist and her performing art evokes a passionate debate in the world of music critics today. It will therefore be a small attempt at showing dialogue between philosophical speculation and living musical experience.

Truth written in notes

A special example of searching for truth in a music score is found in Jerrold Levinson's essay "Truth in Music." Although the author does not express this directly, readers soon find out that they are dealing with the thoughts of an ardent music lover for whom music is the source of a profound aesthetic experience. Levinson seems to declare that in the case of musical enchantment, "feeling and faith" convince him more than cold methods of philosophical analysis. Nostalgia for the metaphysical dimension of music (subtly marked in the title of the book, *Music, Art and Metaphysics*, in which the author included his essay) corresponds with the impeccable logical culture of the argument. Levinson shows readers examples of music which – in his opinion – emanates truth, but that's not all. He also finds examples of insincere and/or false music, which does not mean that it is artistically defective. There is therefore no direct reflection of the truth of music on its artistic value in Levinson's thinking. Music can be hypocritical and at the same time charming, it can also direct the listener's thoughts to banality and obviousness. The most intriguing, however, is the case of music that we consider profound.

The truth of this kind of music dazzles us. It makes us aware that in the experience of beauty there coexists the harmony of good and the necessity of truth, understood in a Platonic way. Levinson clearly suggests the existence of a certain axiological optimum in a work of art, but he does not develop this intuition. Apparently, he counts on his readers being particularly sensitive to musical depth. Unfortunately, he forgets how many authors who write about music do not have that sense (or disposition) that could be called "metaphysical hearing."

Truth revealed in performance

An example of following the other way could be the reflection of Peter Kivy, who does not ask directly about truth, but about authenticity as a part of musical practice. Kivy essentially proposes four notions of authenticity:

1. faithfulness to the composer's performance intentions;
2. faithfulness to the performance practice of the composer's lifetime;
3. faithfulness to the sound of a performance during the composer's lifetime;
4. faithfulness to the performer's own self, original, not derivative or an aping of someone else's way of playing (Kivy, 1995, pp. 6–7).

Kivy's considerations have special value for researchers dealing with authenticity in the performance of early music: Historical truth revealed in musical interpretation plays a considerable role there (Karolina Kolinek-Siechowicz deals with these issues in her doctoral dissertation). Kivy does not offer us any unambiguous solutions. In each of the four cases he mentions, Kivy refers to an enigmatic construct which he calls "an aesthetically correct way of doing the business" or, less ironically, "a viable interpretation."

But what do we do with a situation when an interpretation is convincing for some listeners but not for others? Such a case is described in the brilliant book by Étienne Barilier, *Chinese Piano or Dueling over a Recital* (for the reminder about this book during our conference discussion, I would like to thank Professor Andrzej Hejmej). The novel deals with a dispute between two outstanding music critics over the truth of the "musical message." The debate becomes more dramatic when one of the critics calls the main heroine a fraud, "a crafty little sweetie" (who only pretends to understand the greatest masterpieces of classical Western music), while in the opinion of the other critic the heroine embodies perfect beauty, achieves the heights of mastery and is the greatest virtuoso of modern times. The reader does not know, however, that the fictional name Mei Jin in fact hides a real pianist, Yuja Wang. It turns out that the English-language edition of the book was accompanied by the confession of this contemporary master pianist, who revealed that she was the real inspiration for Barilier. Yuja Wang is one of hundreds of thousands of professional Chinese pianists, but only she has

managed to defeat the legendary Lang Lang. She has become "Steinway's face" and a true concert star playing with the best orchestras and conductors. For several years now, her recitals have drawn crowds of listeners invariably amazed at how someone so fragile and beautiful can play so efficiently, without the slightest trace of fatigue, extracting from the instrument both the most subtle piano and the most powerful forte. She is truly poetic in Schubert, passionate in Rachmaninov, predatory in Bartok. Beautiful gowns, high heels, a model figure and grace – these are just some of the characteristics of the real Mei Jin, whose name is Yuja Wang.

Let's modify Kivy's proposal a bit to construct the criteria for evaluating the Yuja Wang phenomenon. I suggest we leave items (1) and (4) unchanged, and slightly modify items (2) and (3). In the case of Yuja Wang, we are not dealing with historical performance. The revised version of Kivy's proposal would therefore look like this: (2') faithfulness to contemporary performance practice (not "to the performance practice of the composer's lifetime") and (3') faithfulness to the contemporary sound of a performance (not "to the sound of a performance during the composer's lifetime"). The real Yuja Wang is certainly faithful to the composer's intentions recorded in the score. Like her fictional incarnation, Mei Jin, she does not try to show off her pianistic talent and technical skills. When playing Scarlatti, Chopin or Bach, she cares about our aesthetic experience. And yet under her fingers a "modestly wistful fugue becomes heart-rending" (Barilier, 2012, p. 10). It is difficult to find another pianist who would give such contemporary splendor (thanks to the power of his or her own talent) to someone else's work (it is enough to recall the variations on the *Turkish March* or the daring interpretation of *Flight of the Bumblebee* played by the real Yuja Wang). It is difficult for a classical pianist to move with such grace between the roles of a virtuoso and a pop-culture star: She plays her own improvisations, in interviews she quotes Coco Chanel in connection with Gustav Mahler, impeccably performs Rachmaninov's Concerto dressed in a provocative red dress revealing beautiful, long legs and the gently muscled arms of a pianist athlete.

Why then, in the aforementioned book, does one critic see Mei Jin as a mindless seducer who moves around the maze of Chopin's Sonata modulation as if she were in a supermarket on the first day of the sale? What makes it difficult for this critic

to appreciate the pianist is her visual beauty. Mei Jin never looks tired and does not lose control over herself. She never plays a false note. The critic, who reaches the heights of irony in his pejorative assessment of the pianist, suggests that she can be judged positively if and only if you are in love with her. That is how the critic tries to make his colleague's arguments look ridiculous: According to him, in that case, true professionalism precludes being delighted with the talent, virtuosity and, in general, the phenomenon of this young Chinese woman.

Yuja Wang's "truth" is hard to tolerate for listeners accustomed to the seriousness of the concert hall and to musical beauty produced through tears, sweat and suffering. Perhaps it should be listened to from behind a curtain? But in that case, there would be no chance to face the uncomfortable truth why concert halls are bursting at the seams today. In real concert life, Yuja Wang wins, because crowds of people believe her art and want to see what her "truth" sounds like. Probably few listeners realize that the path to Yuja Wang's current lightness and grace led *per aspera ad astra*. At the age of four, she had to practice six hours a day, and at the age of ten – eight hours. Thanks to this, the hands (and the brain) of Mei Jin / Yuja Wang are a bit more "inhuman" than the hands of children who spend only an hour at the piano every day.

In Barilier's book, the fictional Mei Jin also wins, because in the finale of the story the critic-disbeliever suddenly undergoes an internal change. We do not know if he has thought through his colleague's arguments. Maybe he just "heard" and "believed"?

References

Barilier, E. (2012), *Chiński fortepian. Pojedynek wokół pewnego recitalu* (Polish edition of *Chinese Piano or Dueling over a Recital*, transl. J. Giszczak). Warszawa: Noir sur Blanc.

Kivy, P. (1998). *Authenticities. Philosophical Reflections on Musical Performance*. Ithaca: Cornell University Press.

Levinson, J. (2011). *Music, Art and Metaphysics*. New York: Oxford University Press.

Piotr Słodkowski
ORCID 0000-0003-0515-2399
Academy of Fine Arts in Warsaw

Truth of a Painting, Truth of Matter: Robert Rauschenberg, Henryk Streng, and the History of Art

Abstract

Starting from the issue of truth in painting, the text presents a methodological discussion on inter-image relations. The starting point for this thinking is a radical juxtaposition of two artistic practices: Rauschenberg's erasures on a drawing by de Kooning and Henryk Streng scratching out the signature on his own paintings. Both these interferences facilitate an analysis of the transition from approaching images intertextually to a research viewpoint that – quite the opposite – enhances the material aspect of visual representations.

Keywords: art history, modernism, Holocaust, intertextualism, new materialism

Truth in painting or, to put it better, the truth of painting is a convenient category that can be used in describing a large part of art history. The founding fathers of the modern-day version of this discipline traced the truth of a period in its works of art, although they adopted very different optics. Heinrich Wölfflin, founder of formalist art history, studied the roots of style. Outlining general methods of observation and representation, he reached the innovative conclusion that "optical possibilities" restrict the artist's vision and "not everything is possible at all times" (Wölfflin, 1932, p. 11). The second founding father, Erwin Panofsky, in his iconography and iconology project asked about the meaning or the content of a work. The maximalist version of the method, i.e. iconological interpretation, was precisely meant to reveal the truest meaning of a painting and explain it as a symbolic form; a symptom or cultural manifestation that best expresses its time (Panofsky, 1939).

Both of the discipline's 20th-century foundations have been subject to multidirectional criticism, especially from the viewpoint

of the hermeneutics of images and Marxist history of art. Hermeneutics accused iconology of an artificial dualism of form and content, while Marxism blamed stylistic analyses for cultivating the autonomy of a work of art and isolating it from the social context.

But those critics of the *status quo* also met with fundamental criticism, again in the name of the truth of painting. Polemics with their thinking about paintings resulted in one of the most important postmodern debates concerning visual arts. This was Jacques Derrida's "Restitutions," a treatise concluding his famous book *La Vérité en peinture – The Truth in Painting* (1978; cf. Polish edition – Derrida, 2003). As a reminder: The dispute centered around Vincent van Gogh's famous painting *A Pair of Shoes* (1886) and the answer to the question: Who do they belong to? The gravity of the debate, of course, does not lie in the simple identification of the owner but in the clash of completely different ways of thinking about the painting and the truth it reveals.

Martin Heidegger gives his attention to *A Pair of Shoes* in his fundamental text "The Origin of the Work of Art."[1] This is not the place to quote his complex argumentation in full. Suffice it to say that Heidegger incorporates *A Pair of Shoes* into his thoughts on what is thingly in things; on what is the "base of things" and what is their essence (Gr. *Hypokeinenon*) before it is "attacked" by the tradition of thinking organized by oppositions, such as form versus matter. The thingness of a thing can only be captured in its usefulness, hence the philosopher paints a vision of the work of a peasant woman who does not think about her shoes but stands and walks in them, as this is how they actually serve. Why, then, is van Gogh's painting invoked here? Because – to skip to the conclusion – the truth of the work is not the work, and is not manifested in the work: There is a difference between a shoe's "being" and its "existence." Hence the truth or the ontological difference emerges in painting, it is "clad" in painting.

Meyer Schapiro, who ignores the philosophical stake of Heidegger's thinking, attacks him for his carefree approach to paintings, and does so from the viewpoint of Marxist history of art.

---

[1] I outline the arguments of Heidegger and Schapiro, contained respectively in "The Origin of the Work of Art" and "The Still Life as a Personal Object – A Note on Heidegger and Van Gogh" after Derrida's "Restitutions," in: *The Truth in Painting*, transl. G. Bennington, I. McLeod, University of Chicago Press, Chicago 1987, pp. 255–382.

He claims that if the philosopher had bothered to reconstruct the context in which the picture was painted, he would have known that the shoes had nothing to do with the countryside, or with a peasant woman, but must have belonged to van Gogh himself, who was living in Paris at the time he was working on the painting (1886) and, as Schapiro underlines, was "a man of the town and city." Only the social history of art enables us to stay close to the truth about a painting.

Derrida, meanwhile, like a just judge, weighs the arguments of both writers, only to show in the end that they are both wrong. Squeezing the shoes onto a peasant woman's foot in one case, and onto the painter's foot in the other, they in fact project two false hallucinations onto the painting. To Heidegger, the truth revealed in painting is the truth of things (*aletheia*), to Schapiro – a faithful representation referring to an object (*mimesis*). The shoelaces visible in the bottom right-hand corner of the canvas act like a snare trapping these two interpreters. And to whom do the shoes really belong? According to the author of "Restitutions," not to the wearer/owner/signatory but to painting itself, to a multiplying enfilade of paintings. That is how an intertextual game emerges among the works of Jean-François Millet, Nils Kreuger, René Magritte and others.

According to Derrida, painting is only a painting of painting itself [2]. It is believed that the idea that paintings housed in a giant labyrinth gallery only refer us to other paintings was decisive for the deconstruction of the metaphysics of presence. According to Stanisław Czekalski, "we could therefore regard the famous slogan of deconstructivism that 'there is nothing besides text' as a paraphrase of the observation that there is nothing besides a gallery of representations in which paintings do not open passages into the external world of direct presence" (Czekalski, 2006). The truth of painting – in the end – is the truth of iterative repetition, the truth of simulacrum images understood in a virtual sense: traces, quotes, anagrams which have meaning thanks to similarity to other representations.

The legacy of intertextualism turned out to be a refreshing stimulus for historical and artistic research, which made the most mature use of it in the 1990s in Mieke Bal's preposterous

---

[2] I refer to Derrida's *La dissémination*, as characterized by Czekalski, 2006, p. 19.

art history (Bal, 1991, 1999)[3]. On the reverse of this paradigm, however – as Bjørnar Olsen wrote, is the "tyranny of text." Meanwhile, moving beyond disembodied representations and textualism means asking differently about the truth of painting: asking about the truth of matter. How do things remember and what meanings do they materialize (Olsen, 2010)? Here is a reorientation which – as we know – is already visible in the landscape of the new humanities in a wide sense. However, I think it is key to bring up these issues specifically in art history. It is not only because art history works with artefacts first and foremost. At this point, I want to consider the idea that new materialism has turned out to be an especially valuable viewpoint for research on Central European art. To illustrate this, I propose we ask about the truth of matter in the case of paintings by two artists from two geohistorical extremes: American painter, installation artist and pop-art precursor Robert Rauschenberg, and Polish-Jewish modernist Henryk Streng, one generation older, who after the war kept his Holocaust-era identity of Marek Włodarski.

Streng never negated the picture as a cohesive painterly quality and was never interested in any kind of inter-media experiments. Rauschenberg, on the other hand, was just the opposite. He radically broke with the modernist principle of uniform composition that had still been followed by his predecessors, the abstract expressionists from the New York School circle: Willem de Kooning, Jackson Pollock, Mark Rothko. Rauschenberg dissociated himself from them, developing the idea of combine painting: To his canvases he stuck press clippings, advertisements, objects found on the street. This artistic stance is considered to have had its symbolic beginning in a gesture which I suggest we see as being closely related to the practice of Streng. In 1953 the 27-year-old Rauschenberg visited de Kooning, a painter at the height of fame at the time, and asked him for one of his drawings. Upon receiving the penciled sketch, he ostentatiously used an eraser to remove the drawing, leaving only some very faint marks on the paper. This gesture has gone down in history as a manifestation of an artist's own position at the cost of his predecessors, and in fact it can appear as a modern-day version

---

[3] Cf. the excellent discussion on her thinking in Czekalski, 2005, pp. 119–149 and Czekalski, 2006, pp. 222–258.

of the old Vasarian narration according to which the pupil must absorb and surpass the achievements of the master.

One of the foundations of innovative reflection on Central European art is sensitivity to the fact that analogous practices and operations on a work of art have different meanings in different places around the world (cf. Piotrowski, 2011). Comparing the gesture of erasure in the case of Robert Rauschenberg and Henryk Streng is in all likelihood the most distinctive illustration of this methodological conclusion.

Streng's World War II experiences, although extremely valuable as a case study for art history, are consistent with the almost universal biographical experience of Holocaust Survivors who lived under "Aryan papers."[4] Streng was a Polish Jew living in what was Lwów in Poland at the time (now Lviv in Ukraine). After the city was occupied by Nazi German forces (June 1941), he was placed in the Janowska Camp. Streng fled the camp "probably in August 1942," and found refuge thanks to the Polish Council to Aid Jews ("Żegota") established late that year. This was when the artist obtained "Aryan papers," i.e. false ID documents fabricated on the basis of genuine birth certificates, "certificates of people who had really lived and died," in this case a certain Marek Włodarski.

Between 1942 and 1944 when Streng/Włodarski found his way to Warsaw, his drawings and paintings underwent all kinds of interference with their matter, which resulted in the erasure of the old signatures. These occurrences were uncovered by IR (near-infrared) photography, which revealed hidden layers on five of the painter's inter-war canvases; the photographs were taken in 2017 by art conservator Piotr Lisowski at the National Museum in Warsaw[5]. Upon examination, it was found that fragments of the paintings had been interfered with, but not for conservation reasons (retouching of damaged areas or reconstruction of the pictorial tissue). The texture is not continuous in these compositions, and in two cases the interventions are especially radical.

---

[4] See the most extensive calendar of the artist's life in Chrobak, 2010, pp. 111–117. The issues discussed further on are also covered in a large part of my latest book (Słodkowski, 2019).

[5] The paintings in question are: *Composition of Forms* and *Man with a Pipe* (both 1926), *Farewell to the Soldiers* (1933), *Family with a Bouquet* (1936) and *Musicians at a Purple Gate* (1937; the museum call numbers for these paintings are, respectively: MPW 312; MPW 314; MPW 319; MPW 318; MPW 843). The photos were taken at the National Museum on Sept. 14, 2017. My thanks to Piotr Lisowski and the museum's conservator team for their work on the photographs.

In the painting *Man with a Pipe*, the place with the signature was probably sandpapered off, while the photo of *Composition of Forms* reveals expressive traces, vertical lines – micro-grooving that could only have been left by a hard and thin tool, like a paintbrush handle. Similar interventions can also be seen in prewar drawings. On many of them, in place of the signature you can see ragged paper fibers and bleaching that stands out from the yellowing background in other places of the same work. Just like ragged traces of de Kooning's original drawing survived Rauschenberg's interference, we can see bits of the letters "STRENG" showing through where sections were erased by the painter from Lwów.

Bjørnar Olsen's question: "How do things remember?" allows us to consider a dimension of the truth of a painting that is unthinkable in the light of the Derridean enfilade of disembodied works, restitutions, quotes endlessly referring to one another. The truth of a painting as the truth of matter has the irreducible value of a historical source which – specifically – testifies to very different albeit important existential experiences. Erasing a work of high modernism, Rauschenberg entered the scene as a young artist who challenged the "signature" of masterly style and opened art up to postmodern sensitivity: the creation of a picture from elements of the outside, clichéd world, incorporating camp as well as non-normative identities, explaining it outside any model interpretations, in the spirit of deliberate misreading.[6] In a very different way, Streng's interventions also support moving beyond the modernist interpretation of a work and the artist's identity. Because, if the conservative history of Polish modern art implicitly assumed a flat, uniform "Polish" identity of the artists and was founded on understanding a work in terms of pursuing consecutive -isms (avant-garde styles), then the truth of a painting as the truth of matter thwarts this logic. Streng's scraped canvases and pieces of cardboard – as a case study – speak of the biographical experience of a Holocaust Survivor who was a Polish Jew or a Pole of Jewish descent. Examined with an emphasis on their material form – as a "visual concretum" and a theoretical object – they open us up to writing a different history of art[7]: more egalitarian because it transcends the avant-garde idiom; more

---

[6] The opening-up aspects of Rauschenberg's art have been pointed out by authors including Markowska (2012, pp. 14–15).

[7] I elaborate upon the questions briefly outlined here in my latest book (Słodkowski, 2019).

transdisciplinary because it enhances the value of the artefact, which is absolutely essential in multi-area research like studies on the Holocaust; more agency-focused because it is especially sensitive to agentivity and the identity-based roots of non-Western artists. The history of Central European art starting from the 1980s owes a great deal to critical and postcolonial interventions that made us sensitive to the epistemology of difference. Today it is new materialism that could come to our aid.

References:

Bal, M. (1991). *Reading "Rembrandt": Beyond the Word–Image Opposition*. Cambridge: Cambridge University Press.

Bal, M. (1999). *Quoting Caravaggio: Contemporary Art, Preposterous History*. Chicago: University of Chicago Press.

Chrobak, J. (2010). Henryk Streng / Marek Włodarski. Kalendarium. In: J. Chrobak et al. (Eds.), *Czarownik przy zielonej skale. Henryk Streng / Marek Włodarski* [Sorcerer at a Green Rock. Henryk Streng / Marek Włodarski] (pp. 111–117). Poznań: Galeria Piekary.

Czekalski, S. (2005). Semiotyka widzenia i preposteryjna historia obrazów Mieke Bal [Mieke Bal's semiotics of vision and preposterous history of paintings]. In: M. Poprzęcka (Ed.), *Obraz zapośredniczony* [Intermediated Picture] (pp. 119–135). Warszawa: Stowarzyszenie Historyków Sztuki.

Czekalski, S. (2006). *Intertekstualność i malarstwo. Problemy badań nad związkami międzyobrazowymi* [Intertextuality and Painting. Problems of Studies on Inter-image Relationships]. Poznań: Wydawnictwo Naukowe UAM.

Derrida, J. (2003). *Prawda w malarstwie* (Polish edition of *The Truth in Painting*, transl. M. Kwietniewska). Gdańsk: słowo / obraz terytoria.

Markowska, A. (2012). *Dwa przełomy. Sztuka polska po 1955 i 1989 roku* [Two Watersheds. Polish Art after 1955 and after 1989]. Toruń: Wydawnictwo Naukowe UMK.

Olsen, B. (2010) *In Defence of Things: Archaeology and the Ontology of Objects*. Lanham: Altamira Press.

Panofsky, E. (1939) *Studies in Iconology: A Study of Humanistic Themes in the Art of the Renaissance*. Oxford: Oxford University Press.

Piotrowski, P. (2011) *In the Shadow of Yalta: Art and the Avant-Garde in Eastern Europe, 1945–1989*. London: Reaktion Books.

Słodkowski, P. (2019). *Modernizm żydowsko-polski. Henryk Streng / Marek Włodarski a historia sztuki* [Jewish-Polish Modernism. Henryk Streng / Marek Włodarski and the History of Art]. Warszawa: Instytut Badań Literackich PAN.

Wölfflin, H. (1932). *Principles of Art History. The Problem of the Development of Style in Later Art*, transl. M. D. Hottinger. New York: Dover Publications.

Maciej Junkiert
ORCID 0000-0002-7415-4219
Adam Mickiewicz University, Poznań

The Polish History of Literature as a *Lieu de Mémoire*

Abstract

The paper contains an analysis of the genealogy of 19th-century Polish research in the field of literary history. My inquiry contains a comparison between literary research in Germany and in Poland. From this point of view, literary history was an important factor in the process of building a modern nation. Furthermore, literary historians also played the role of indispensable authorities on the cultural past and present: in Germany as professors of universities, and in Poland during the partitions as intellectuals and writers.

Keywords: history of literature (discipline), Adam Mickiewicz, Maria Janion

On August 9, 1789, a young German aristocrat took a stroll around Paris where he had been staying for but a week. His name was Wilhelm von Humboldt[1] and he was an aspiring writer and philosopher. Having completed a semester of law school in Göttingen, he had decided to visit a place that was attractive to all writers at the time – the capital of a France in revolt. That day, he made his way to the Bastille, which only three weeks earlier had been a fortress and a prison. Since the Middle Ages, it had ruled over the city and, although with time it lost its former military function, it was still a place which embodied the royal reign and domination. Now, after the fortress had been taken and plundered, and its prisoners released, it was the first tangible testimony to the power of the revolution. In this case, the victory came with a desire to completely destroy and wipe this

---

[1] The literature on Humboldt is quite extensive and in recent years, due to several important anniversaries of his multifaceted political activity and research output, a number of new books have been published. See the essential intellectual biographies (T. Borsche, 1990; F. C. Beiser, 2011, ch.1; Quillien, 2015).

place off the map of Paris. During Humboldt's stay, the Bastille was being demolished, and the prison could only be visited on Sundays, which the traveler was happy to do.

Although the German aristocrat's visit took place at an extraordinary time, there was scant trace of the revolutionary events in his travel log from France and his stay in the capital. The young man visited parks, palaces and churches; he admired works of art, observed the monumental propaganda of Parisian architecture, and paid no attention to politics and current affairs. However, his walk around the Bastille inspired him to take a closer look at the motives of those who had challenged the fortress. Humboldt wandered around the walls that symbolized royal tyranny like no other place did. He saw the victors' enthusiasm confronted with the vestiges of suffering endured by many a generation of prisoners. In his view, an unarmed crowd lunging at a great fortress must have had a weapon that the royal soldiers lacked. It was despair that gave the masses an unstoppable energy, triggered by the threat that the emerging transformations might be stifled (Humboldt, 1916, pp. 119–121).

Humboldt only saw what he wanted to see and what matched his worldview. He ignored the Bastille's demoralized garrison, who were not going to lay down their lives for the compromised king. Also, that defenseless crowd which attacked the fortress actually had 20,000 muskets at their disposal as well as a number of cannons obtained earlier from Palais des Invalides. Furthermore, the attack was launched by the National Guard commanded by La Fayette, hence the notion that it was an unorganized entropic force which took over the fortress is untrue (Furet, 2008).

Humboldt described the Bastille as an example of a place which had lasted for centuries, enjoying its grim fame and devouring countless lives of condemned men, only to collapse as a result of an impulse, a fake rumor that the armies loyal to the king had planned an anti-revolutionary attack on Paris. As a writer fascinated by the Enlightenment's philosophy, particularly its British version, he both admired the authenticity of the freedom that the nation emerging before his eyes represented and feared the disintegration of the familiar world of traditional sociopolitical relations. Along with a number of representatives of German culture at the time, Humboldt asked himself the following question: How can similar goals be achieved without inciting violence and overthrowing existing systemic institutions?

Just 20 years later the experiences and observations from revolutionary Paris would come in handy to Humboldt. After Prussia's defeat to Napoleon, Humboldt was dismissed from his ambassadorship to the Vatican and appointed director of the Department of Education and Religion. His task was to reform secondary education and create a new type of university in Berlin.[2] This story about Humboldt is relevant to me because, as a historian of literature, I have been studying the emergence and development of my discipline. No one made a greater mark on its development than this German politician and philosopher. This stems from the fact that when he was creating individual departments and hiring faculty, Humboldt defined the history of literature in a particular way. From that moment on, it would aid the process of educating society and shaping its opinions.

In the first decade of the 19th century the history of literature exhibited a surprising lack of interest in facts and texts from the past. While the goal of classic historiography was searching for knowledge about the past and its reconstruction, the history of literature was given other tasks. In the view of August Wilhelm Schlegel, one of the patrons of historico-literary research, the military defeat of Prussia necessitated an intellectual response in the form of evoking ancient virtues, the notions of freedom and glory of the old Germanic people. His goal was to confront contemporary citizens of the German states with a constructed historical vision so their current condition could find counterbalance in the heroic message from the past, particularly from old myths, tales and poetry. The history of literature framed in such a way was to become the glue of the newly conceived national community. It was implied that a lack of commendable results or achievements from the past would never hinder the effectiveness of the prehistoric authority. From the very beginning, the

---

[2] A Humboldt university was referred to as a certain type of institution simultaneously conducting research and educating students. For around 20 years now, it has not been discussed in these terms any more, except in Poland. Critics have consistently stressed that the theoretical assumptions of Humboldt's ideas were never fully implemented, and soon after his short reformatory mission was over, his conservative successors quickly made all schools subject to the state, against Humboldt's premise of a university's partial independence and freedom of research and education. Years back, Bill Readings wrote an insightful book (Readings, 1999) about the impact of Humboldt's model on the development of the humanities.

first Polish historians of literature were fascinated by the German models, and traces of this historico-literary ideal, rooted in the politics and education of the time, are scattered across the discipline's 200-year history.

It is my assumption that the Polish humanities have never viewed the history of literature as one of many disciplines. I believe that Maria Janion diagnoses this issue most effectively in her research output, starting from her doctoral dissertation on the early literary career of Zygmunt Krasiński, all the way to *Niesamowita Słowiańszczyzna* [The Incredible Slavs]. In her writing, historico-literary research had always been synonymous with a special bond between tradition and the cultural past. That is, until the end of the past decade, when Janion lost her faith in the traditionally viewed connection with the past:

> Today the debate about universal and vernacular values, which has been around since the 18th century, has come to a dead end. This is a result of the fact that the ability to read old Polish culture is vanishing, and instead, it is treated as little more than a collection of ideological quotations (Janion, 2006, p. 7).

It is not my intention to deprecate the works of a great many prominent researchers, among them Zygmunt Łempicki, Henryk Markiewicz, Kazimierz Wyka, Stefan Sawicki, Jerzy Ziomek, Teresa Walas, Ryszard Nycz or Przemysław Czapliński (to name just a few), who have made a significant contribution to pre- and postwar historico-literary reflection. However, Janion's approach to the history of literature has always been special. It was based on highlighting the birth of historicism, which was fundamental for the development of European culture. She has always stressed that if it had not been for Romantic historicism, which arrived in Poland from France and Germany, the Polish identity would look quite different. It was almost impossible to imagine Polish culture without Romanticism, while understanding this culture without the role of historians of literature seemed unfathomable.

Has the history of literature, framed in such a way, become a Polish site of memory?[3] I was inspired to ask this question by

---

[3] Research on memory has already entered the phase where textbooks, scripts and monographs aggregating the current state of knowledge are written. Several of these most recent publications offer a sound overview of the current state of memory studies: Erll, Nünning, 2010; Gudehus, Eichenberg, Welzer, 2010; Erll, 2011; Feindt, Krawatzek, Mehler, Pestel, Trimçev, 2014.

a reflection of François Hartog, who recently posed a similar question regarding history as a European site of memory. Hartog follows up and elaborates on the reflections of Pierre Nora, who supported his concept of sites of memory (which he has been developing since the early 1980s) with the loss of a social and intellectual bond with the past. Nora's concept – contrary to other groundbreaking views offered by Halbwachs, Le Goff, Ricoeur, or Jan and Aleida Assmanns – was based on the belief that the experience of modernity is characterized by a sense of accelerated time, the need to constantly document and record it, and to construe institutionally aided memory which will be responsible for a relationship with the past (Nora, 1996).

So as not to repeat Nora's arguments, which are quite well-known, I would only like to note that according to him, all the fields and disciplines that are affected by a similar crisis, will struggle. The history of literature is facing a crisis whose focal point, as Nora believes, is the discussion around the discipline's scientific birth. Referencing a seven-volume dictionary of French sites of memory, Hartog develops his own idea about two systems (or regimes) of historicism. One of them is focused on the past, the other on the present, whereas the year 1789 is treated as a symbolic shift in the paradigm (Hartog, 2015). According to Hartog, history travels a long way toward commemoration. Its task is to help (cultural and collective) memory bring back from oblivion what has shaped our contemporary condition. Hence, it is worth asking the question: What tasks should we set the history of literature? What should it commemorate and where is its place?

If we concur that the Polish history of literature has become a site of memory which requires permanent and institutional support because it has lost the vestiges of its former social legitimization, we would also have to admit that the effects of this process are visible to the naked eye. Undeniably, my interest in this subject is augmented by the fact that in the last decade, there has been a steady decline in the number of Polish scholars and linguists who refer to themselves as historians of literature, and to their profession as the history of literature. The number of historico-literary doctoral dissertations and post-doctoral degrees has also diminished. For quite some time now, Polish literary studies have been moving away from their historico-literary roots. Furthermore, this dispute, which has lasted for at least three decades, directly affects the scope of contemporary

methodological research and it will surely determine the future of the entire discipline.

Sonderweg, or evolution
without a revolution

The outbreak of the French Revolution was cheered by the German-speaking states, particularly by the circles of writers, scientists and thinkers. The evolution of their views, especially the aforementioned Humboldt, Schiller, Fichte and Friedrich Schlegel, constitutes an important research area in German intellectual history (Saine, 1988). What made the great enthusiasts of revolution instantly change their approach toward the events evolving in Paris? In one of her recent books, Rebecca Comay interprets it as a certain perceptual error. For a long time German thinkers did not approach the revolution in France as a universal phenomenon but only as a local issue of the French (Comay, 2011). Complications arose when revolutionary France started viewing the German lands across the Rhine as a space open for expansion whereas Prussia and Austria were now its mortal enemies. In 1807 in Berlin, which had been taken by the army of the victorious Napoleon, J. G. Fichte gave the famous *Addresses to the German Nation* in which he redefined the role of the past in shaping the future identity of all the Germans (Borchmeyer, 2017). It was Fichte who greatly inspired Humboldt when the latter was founding the University of Berlin, placing emphasis on history, philosophy, language and literature as the cornerstones of education and key research areas.

At that point, the history of literature did not become a secondary discipline. From the beginning, it competed with other disciplines for influence, positions and level of funding. German philology, which aggregated literature and language studies, was soon scornfully tagged as *Brotwissenschaft* – a science for money (bread), which made it stand apart from more noble and older faculties (Meves, 1994). However, this younger discipline proved to be a fierce competitor. Due to the tasks involved in educating teaching staff for the growing number of schools, it could count on increasing support from the Prussian government, and as the Humboldt model of the university became more popular, from other German states as well.

History of literature
as a national mission

I would like to discuss this issue without moving beyond the first half of the 19th century, and, to be precise, by focusing on selected works of two people who were particularly vital to Polish culture. They are Kazimierz Brodziński and Adam Mickiewicz.[4] Both of them, and each in his own way, initiated a certain type of historico-literary reflection in Poland. They were not the first Polish historians of literature. Mickiewicz himself would most likely strongly oppose being called one. Nevertheless, considering the nature of their writing and scholarly activity, they provide two turning points in the first half of the 19th century. Between 1822 and 1830, Brodziński taught Polish literature at the University of Warsaw. Mickiewicz, on the other hand, taught Slavic literatures in Paris between 1840 and 1844. Their activity became a reference point for a number, if not the majority, of historico-literary efforts in Poland, and not only from the era of the partitions.

Brodziński

In 1818 Brodziński wrote in his critical treatise on Classicism and Romanticism: "Poetry is a mirror to any century and nation." However, if literature reflects and represents the truth about a nation's past, then only a historian of literature is able to interpret that reflection and verbalize that truth. This famous article may be analyzed simultaneously with the critique of the myths of historicism by Nora. He treats the mirror reflection as one of the primary metaphors illustrating the tension between memory and history. A historian of literature who uses the image of a mirror in evoking and explaining the past is suggesting that they are able to show the beginnings and evolution of a given phenomenon as well as its role in the life of the national

---

[4] The bibliography on Brodziński and Mickiewicz is beyond extensive; however, as far as the historico-literary dimensions of their activity go, the sources are much more scarce. Works by the following authors are noteworthy: Rościsław Skręt, Eugeniusz Klin, Wiktor Weintraub, Maria Prussak, Marta Piwińska, Piotr Śniedziewski, and Michał Kuziak (among others). Few works are written on this topic, and their greatest shortcoming is weak placement of the Polish historico-literary narratives against the French and German models.

community (in the case of Brodziński, this phenomenon was the idyll). However, in reality, the only thing that can be shown is a series of differences between the past and the present. The questions who we were in the past and why we are no longer this way cannot be answered.

A little later (at that point, he was a professor at the University of Warsaw), Brodziński wrote the following:

> If a nation's insignificance, futility of efforts, frenzy and misfortunes leave a sorrowful impression over the course of history, the history of a human society's enlightenment always brings comfort. It is that history where we see the most Providence, which prods people toward greater and greater dignity. Temporary frenzy and suffering become irrelevant in the course of time. Yes, a watchful eye perceives it as necessary repose, necessary disorders that lead to an even greater rebirth. Frenzy and affectation fade away, much like a redundant flower, but each grand idea lasts for centuries and gives birth to another one. Each truth that has been revealed, even if obscured for a moment by clouds, lives on, like the sun, shines even more brightly. All the hindrances and altercations only serve it well and increase its triumph. Only he who looks beyond the present can understand vocation, dignity and the meaning of life, and only he who firmly believes in reaching for ever greater glory and happiness of people can appropriately comprehend literature. This zealous belief in it is the first spring toward loving it, the first stance which allows for its appreciation (Brodziński, 1872, p. 100).

Brodziński's opinions trigger two observations. Firstly, his beliefs are a contamination of the views of key German thinkers, particularly Winckelmann, Herder, Schiller and Friedrich Schlegel. Secondly, his concept of the history of literature has been imported entirely from the German background. Thirdly, there is one fundamental difference, because Brodziński formulated his vision of the institutionalized history of literature for a nation without statehood. Finally, Brodziński hides his role of author and acting subject of the historico-literary narrative, invoking the authority of literature, to which he only loans his voice. Having synthesized the features of early Polish culture, Brodziński moved on to clarify the purpose of his reasoning and the research method:

> These are the general properties of our old literature, which I will try to demonstrate over this course. It shall not be interesting or engrossing, though. The entire history of our national education has

been a tedious road, and many of its sections still need paving. At times, it will take us through enjoyable and pleasant locations, but often we will have to wade through drifts, arid and barren lands, and, even more often, we will face obstacles. However, the destination where I wish to take you, Gentlemen, should make this journey more enjoyable. This goal is to serve justice to our ancestors, and keep their legacy alive (Brodziński, 1872, pp. 112–113).

In my view, Brodziński's concept and the aforementioned words by Janion refer to the same imagined historico-literary community. While Brodziński was one of its first exponents, Janion is undoubtedly one of its last representatives. In the spirit of this agreement, national culture requires special guardians and depositories of its values who will be capable of recalling and explaining what others have forgotten or what they have never learned. The history of literature understood in this way has never detached from its Romantic roots, and historians themselves never doubted the importance of their mission.

Mickiewicz

If we wanted to apply D. R. Kelley's (Kelley, 2010) formulation on two different types of historiography, Brodziński would definitely represent the history of Thucydides whereas Mickiewicz would surely be a follower of Herodotus' mission. Brodziński believed in the scholarly and educational dimension of historico-literary research, in knowledge derived from the correct reading of a message from the past. Mickiewicz, on the other hand, taught Slavic literature in its cultural and political context, which rejected the existence of the one and only true interpretation. Its scholarly dimension was dismissed and excluded by him at the very beginning, in the preface to the German edition of his lecture:

> As I have been for the most part unable to use historical documents, I had to begin with the only means that were at my disposal: my memories. Whatever I felt or noticed during my stays in various Slavic countries, what I recorded in my mind from my old works on history and literature, and particularly what I contracted from the spirit which is animating nations these days – this is all I had. And this is what I shared with my listeners. The literature course at Collège de France is aimed more at presenting the results achieved by the discipline than a detailed analysis. Among the attendees of Collège de France, there are those who know the details as well as

the professor and who cannot be taught as students. My audience was largely made up of Slavs. All these factors had a great impact on the shape of my lectures.

Whenever I was going to speak to my listeners, I stood in front of them without a prepared speech, often without any written notes. The raised topic often took me right to the core of related literary and philosophical questions, and by improvising, I presented the results of my old works as well as my deepest personal feelings (Mickiewicz, 1997, p. 9).

Brodziński's intention was to raise and shape a national sense of greatness, morality and community by showing the distant past of Polish literature. Mickiewicz taught literature for political reasons, at a time when a sense of impending conflict between the West and the East was on the rise (pardon the simplification). Brodziński's history of literature was timeless and universal while Mickiewicz's was engaged and temporary. Brodziński studied sources and reconstructed the literary past according to the principles of philological and historical critique. Mickiewicz ostentatiously rejected the study of literature, claiming – contrary to the facts and pedagogical empiricism – that he drew on his memories and readings of his youth. Although he was reluctant to admit it, Brodziński's formulation of thoughts and his historico-literary narrative were often inspired by German philology, aesthetics and emerging history of literature. Mickiewicz, for whom various German publications (still not fully described by the lecture's researchers) were also a key source of knowledge about Slavic culture and which he was just as reluctant to quote, ostentatiously rejected the approach of German researchers toward literature and culture. He confronted it with Slavic culture – to use Larry Wolff's term, which is quite apt in this context – as "the invented East,"[5] enveloped in a mist of researchers' and readers' ignorance supplanted by averse civilizational and political fabrications, fictions and lies.

If I were to point to a figure who unifies Brodziński's and Mickiewicz's efforts, and who is permanently inscribed in the Polish historico-literary narrative resurfacing in research and

---

[5] Wolff's incredibly important and revealing book (Wolff, 1994) has not yet been appropriately received or followed in Poland, although thanks to the evolving research on the postcolonial dimension of Polish culture in the 19th century, this state is likely to change before too long. See (among other sources) Kuziak, Nawrocki, 2017.

published books, it would be Guślarz [Wiseman, a Slavic shaman-like figure] during the ritual of Dziady (Forefathers' Eve). He is anchored in the past but also leans into the future, a future in which he aims to play a part. In Polish historico-literary research, the goal has always been to get in touch with the ancestors and do them justice, because without this contact it was (and still is) impossible to define the Polish identity.

The literary historian used to be an elitist, niche profession, one designed to serve the community. He or she also deeply believed that there would always be a place for their research in the area of social interest. They tried to transfer knowledge about the past that could fill in the blanks but also help build a link between bygone times and the broadly defined experience of the present. Historians of literature are struggling to adjust to the current evolution of the humanities and the changing role of the university. Their Polish-centered approach, which used to be a strong suit, now has become their weakness. We are clearly standing at a crossroads, searching for a new model of research, a new ethos and a place on the map of the humanities.

I am not going to put forward the hypothesis that the beginnings of the history of literature determined its later condition. It would mean embracing one of the most dangerous historiographic myths. Nevertheless, I suspect that the numerous publications pushing one to rethink the history and output of German (Ulrich Muhlack, Hans Ulrich Gumbrecht, Marcel Lepper, Peter-André Alt), French (Pascale Hummel) or British (Simon Goldhill, M. A. R. Habib, Joshua Billings) historiography might inspire contemporary Polish literary scholars to do the same.

<div style="text-align:right">Translated by Katarzyna Szuster</div>

References

Beiser, F. C. (2011). *The German Historicist Tradition*. Oxford: Oxford University Press.

Borchmeyer, D. (2017). *Was ist Deutsch? Die Suche einer Nation nach sich selbst*. Berlin: Rowohlt.

Borsche, T. (1990). *Wilhelm v. Humboldt*. München: C.H.Beck.

Brodziński, K. (1872). Literatura polska [Polish literature]. In: K. Brodziński, *Pisma* [Works], vol. 3. Poznań: Gebethner i Wolff.

Comay, R. (2011). *Mourning Sickness. Hegel and the French Revolution*. Stanford: Stanford University Press.

Gudehus, Ch., Eichenberg, A., & Welzer, H. (2010). *Gedächtnis und Erinnerung. Ein interdisziplinäres Handbuch*. Stuttgart: Metzler.

Erll, A., Nünning, A. (2010). *A Companion to Cultural Memory Studies*. Berlin: De Gruyter.

Erll, A. (2011). *Kollektives Gedächtnis und Erinnerungskulturen*. Stuttgart: Metzler.

Feindt, G., Krawatzek, F., Mehler, D., Pestel, F., & Trimçev, R. (2014). *Europäische Erinnerung als verflochtene Erinnerung. Vielstimmige und vielschichtige Vergangenheitsdeutungen jenseits der Nation*. Göttingen: Vandenhoeck&Ruprecht.

Furet, F. (2008). *Revolutionary France 1770–1880*, transl. A. Nevill. Oxford: Blackwell.

Hartog, F. (2015). *Regimes of Historicity. Presentism and Experiences of Time*, transl. S. Brown. New York: Columbia University Press.

Humboldt, W. von (1916). Tagebuch der Reise nach Paris und der Schweiz 1789. In: W. von Humboldt, A. Leitzmann (Eds.), *Gesammelte Schriften*, vol. 14: *Tagebücher. Erster Band. 1788–1798* (pp. 76–236). Berlin: B. Behr's Verlag.

Janion, M. (2006). *Niesamowita Słowiańszczyzna. Fantazmaty literatury* [The Incredible Slavs: Phantasms of Literature]. Kraków: Wydawnictwo Literackie.

Kelley, D. R. (2010). *Oblicza historii. Badanie przeszłości od Herodota do Herdera* (Polish edition of *Faces of History: Historical Inquiry from Herodotus to Herder*, transl. M. Tomaszewski). Warszawa: Wydawnictwo Naukowe PWN.

Kuziak, M., Nawrocki, B. (2017). *Romantyzm środkowoeuropejski w kontekście postkolonialnym* [Central European Romanticism in a Postcolonial Context], part I. Warszawa: Wydawnictwo Instytutu Badań Literackich PAN.

Meves, U. (1994). Zum Institutionalisierungsprozeß der Deutschen Philologie: Die Periode der Lehrstuhlerrichtung (von ca. 1810 bis zum Ende der 60er Jahre des 19. Jahrhunderts). In: J. Fohrmann, W. Voßkamp (Eds.), *Wissenschaftsgeschichte der Germanistik im 19. Jahrhundert* (pp. 25–47). Stuttgart: Metzler.

Mickiewicz, A. (1997). Literatura słowiańska. Kurs pierwszy [Slavic Literature: First Course]. In: A. Mickiewicz, *Dzieła* [Works], vol. 8, J. Maślanka (Ed.). Warszawa: Czytelnik.

Nora, P. (1996). Between memory and history. In: P. Nora (Ed.), *Realms of Memory. The Construction of the French Past*, vol. 1: *Conflicts and Divisions*, transl. A. Goldhammer (pp. 1–20). New York: Columbia University Press.

Readings, B. (1999). *The University in Ruins*. Cambridge, MA: Harvard University Press.

Quillien, J. (2015). *L'antropologie philosophique de Wilhelm von Humboldt*. Lille: Presses Universitaires du Septentrion.

Saine, Th. P. (1988). *Black Bread – White Bread: German Intellectuals and the French Revolution*. Rochester: Camden House.

Wolff, L. (1994). *Inventing Eastern Europe: The Map of Civilization on the Mind of the Enlightenment*. Stanford: Stanford University Press.

Jan Kutnik
ORCID 0000-0002-4476-7161
John Paul II Catholic University of Lublin
»Artes Liberales« Academy

Truth of the Place and Truth of the Exhibition: "Case Study" of the State Museum at Majdanek

Abstract

The author discusses the problem of truth in museum exhibitions invoking the tragedy of World War II. In this respect, he is very appreciative of the value of Majdanek, Auschwitz-Birkenau, Stutthof and the other museums set up within the compounds of former Nazi German camps. Based on his own research from 2016–2019, he finds that shaping the relations between the truth of the place and the truth of the exhibition, skillfully using authenticity, and finding new forms of "speaking" to visitors are enormous challenges facing exhibitors.

Keywords: Majdanek, Holocaust, museum, truth of place, exhibition

As tangible testimony to the truth of genocide, every site of the Holocaust and every museum dedicated to the victims is unique, crucially important and should be heavily protected. Obviously, this is a completely banal statement with which everyone will probably agree, even people pointing to the inexpressibility of the truth about the atrocities commited against those persecuted by the Third Reich. In the space of the German camps, museums are among the basic instruments of education and are one of the main means of expression and sources of knowledge about the crimes of the Nazis. The history of extermination is first and foremost the story of specific victims, expressed in personal testimonies. Of course there is also a certain amount of facts that are much more than individual, as they describe the fate of prisoners as a group.

The testimony of prisoners, reflecting their personal experience and "the climate" of life in the camp, is subjective. At the level

of objective facts, if only such data is available for processing, we may e.g. "learn" how many people were imprisoned in a barrack and calculate on this basis how much space was allocated per prisoner. Only such a subjective-objective view combined into one narrative – exposure discourse, can bring fulfillment of the "truth" of the place. Statistical facts are part of the truth of a particular place and they complement prisoners' accounts. Obviously, this truthfulness is not diminished by the fact that more precise numbers are not known and probably will never be known. The museum as an institution must carry out a particular function. Above all, it is a way of commemorating the victims and "an eternal witness." Only after that does it fulfill other, no less important roles, such as educator, advocate and exponent of historical truth (or at least knowledge compatible with current findings). This is the ideal that should guide it, and the authenticity of the place should reinforce the desired effect (cf. Unger, 1988; Webber, 2003; Knigge, 2003; Kranz, 2003).

Konzentrationslager Lublin was neither the first camp, nor the largest, nor a model camp, nor the longest operating one. The museum established there was the world's first such institution and site of remembrance created on the grounds of a German concentration camp (which shortly also became a death camp). Indeed KL Lublin was the first Nazi concentration camp to be liberated by the Allies. In this context I would like to treat the museum established there as a special *case study* through which the crucial problems of similar institutions can be described. However, by using the term *case study*, I do not suggest that the analyzed "case" should be considered "afflicted" or an abnormal "clinical case."

The museums created at Majdanek (1944) and Auschwitz (1947) were simply pioneering in Europe and their undeniable value as "symptomatic cases" reveals itself precisely in the fact that they set certain patterns and standards, at the same time providing a point of reference for others. From the quite long history of the Museum at Majdanek I will choose only a few examples/motifs in which one can clearly see various difficulties with maintaining truthfulness and dealing with falsehood in relation to shaping the space and creating an exhibition narrative, as well as their impact on visitors. As time went by, the Auschwitz-Birkenau museum took the initiative in a peculiar rivalry of memory, and Majdanek (Wóycicka, 2009; Cebulski, 2016), getting left

behind, was a clear example that the authenticity of a place was not always at the forefront when faced with the pragmatism of history-based politics (Banach, 2014; Olesiuk, 2011).

1. Truth of the place –
identifying the problem

The question of the truthfulness or authenticity of a place, of all the artefacts or relics of the camps, is not only a theoretical problem of museum studies but also a very practical one. It is vital in the context of various aspects of a museum's functioning, from the arrangement of the space, through the conservation of architectural structures and smaller objects, to the reconstruction of buildings (Knigge, 2003; Webber, 2003; Rymaszewski, 2000, 2003; Ziębińska-Witek, 2011). When time does not work in the favor of living memory nor the material's durability, it is urgent to answer the following question: What action can one take to preserve and present "the truth of the place," and what actions should be considered unacceptable? This is a dilemma that resounds in the background of every single decision of museum curators about any intervention in the fabric of a protected site. The theme of the impact of the authentic "aura" of a given site can often be found not only in theoretical analyses and deliberations but is also clearly present in conversations with both employees and visitors (for more about research on visitors in the context of the importance of the authenticity of a place and site, see Nowacki, 2015).

It is worth emphasizing that these issues seem to be particularly important today, due to new trends and a new situation that has been a challenge for exhibitors since the beginning of this century. Employees of genocide museums and curators responsible for creating the exhibitions face challenges they have never faced before at all or on such a scale. On one hand, we have such events and situations as selfies with crematorium furnaces (KAD, 2008), a camp guard tower as photo wallpaper (Szlachetka, 2015), or a Palestinian-Israeli tour group of teenage boys interested mainly in smoking a cigarette in some nice place within the museum (this is an example from an ongoing visitor study at Majdanek, Kutnik, 2016–2019). On the other hand, the last generation of victim-witnesses is dying; though not numerous, in exceptional

situations they act as guides at sites of memory and they give important personal testimony of Nazi crimes (e.g. Tomasz "Toivi" Blatt, the last prisoner and insurgent from Sobibór, died in October 2015). For institutions talking about World War II genocide, all such people are undoubtedly extremely important. They are practically irreplaceable, above all because they complete the triad of authentic place, exposition and witnessing participant, i.e. "a real thing, place and person" (Moore, 2000, p. 135). When witnesses can no longer testify, who else can be a spokesman for the truth if not a museum, institutionally binding the two elements that are left (see Kranz, 2011)?

In the context of the progressing historical and social transformations mentioned above, the problem can be presented in the form of a few simple questions. How should we exhibit in order to speak (effectively) to future generations? Is the authenticity (truthfulness, reality) of place, artefacts and exhibits still sufficient (to influence the museum's audience)? Is there, perhaps, "too little of" this authenticity? When is authenticity legible in itself, and when does it need to be expounded and how much? Another example from the aforementioned ongoing research (right after the Palestinian-Israeli boys' trip) is the visit of a four-person family from one of the Far East countries who came to Poland for a few days, and in Lublin planned to visit Majdanek (Kutnik, 2016–2019). This was a middle-aged married couple with two daughters, a teenager and a little girl. After about ten minutes of a guided tour providing basic information about the history of the camp, the woman interrupted by suddenly asking, "What was it all about with these Jews?" It cannot be denied the question caused some consternation even in an experienced guide. The visitors were well educated, worked in the IT sector and showed a lot of commitment and interest. As it soon turned out, they knew practically nothing about the historical context concerning the war in Europe and the Holocaust.

In the proposed "case study," looking at the museum I would like to start with thoughts expressed by visitors and collected during the research project (carried out in 2016–2019). The project focused on the conditions of the reception of the space and exhibition at Majdanek. One of the initial stages involved active observation of groups of visitors and the guides at the memorial site. First, the main object of investigation was the coherence of the presented narrative and the elements to which visitors paid

special attention (I wrote down the questions they asked and the answers they received). Additionally, there was a questionnaire that included a request for a free comment on their perception of the museum. One quite widespread view can be found in a literal quote from one of the visitors: "This place should be more realistic. At least one barrack should be arranged as it was during the camp's operation, i.e. bunk beds should stand as they did then and in the same number, etc." (by the way, it needs noting that such an exhibition is actually on display) (ibid.). The simple demand contained in this clear statement can be paraphrased as a kind of imperative. It would express the essence of the functioning of a history museum, being a call to preserve the compatibility or coherence of the museum's exhibition (performance) with the "historical reality" of a given place. Thus we end up with an analogy of the classical definition of truth, where a museum exhibition is the equivalent of "thought" or "judgment" which should "agree with reality," i.e. facts concerning a given place, phenomenon or event (Unger, 1988, p. 18).

2. Truth and falsehood – examples
of encountered difficulties

In order to reflect on the "truth of the place," it is important and intriguing to consider the history of the postwar struggle to preserve the authenticity of the former German camps, and to analyze the factors that influenced the shaping of the first exhibitions (Wóycicka, 2009). The list of problems the management of this institution encountered over the decades of its existence is quite extensive (Banach, 2014, 2016) and includes, among other things, reducing of the area, theft of property (left by prisoners and making up the camp), destruction of material artefacts and objects, various concepts of land development (such as afforestation) and various contemporary modifications in the presentation (being to some extent the result of visitors' behavior). Therefore the current state of affairs is the result of many compromises, but also of old, misguided or unrealistic ideas for arranging the site (Olesiuk, 2011).

In the first years of the institution's operation, the creators of the Museum at Majdanek felt the burden of the place's propaganda potential. On one hand, it provided some protection against the

devastation of objects and degradation of the camp grounds. On the other, it "clashed" on certain levels with concern for historical truth. Posing the provocative question whether the Majdanek museum's exhibition falsifies history, we should answer: If the condition *sine qua non* of falsification is the intention to mislead the viewer or to present a falsehood, then today the answer would be no. Has the exhibition ever falsified history? Here, the answer would be, unfortunately – yes. It happened when the propaganda effect and shaping "the appropriate" social attitudes had much greater value than expressing the truth.

Zofia Wóycicka (2009) presented an excellent analysis of this phenomenon in relation to all museums at concentration and death camps. This manipulation of the truth seems to be particularly glaring with regard to instrumentally treated facts concerning the extermination of the Jews (Banach, 2013, 2016). In the first period of the museum's operation at the former KL Lublin, information about this issue tended to appear and disappear in turn from the exhibitions and the official narrative. Of course, in the clash with the state apparatus and the communist party's current line in a given time, reliability and fundamental standards of presenting facts could not be maintained successfully. The only thing that could be done was efforts to minimize the damage. In practice, it was only after 1989 that the times became pretty calm, without any urgent need for "propaganda" falsity. The last three decades have been a time of spectacular successes in "putting straight" or expanding well-founded knowledge about Majdanek. One example of some extremely important research would be the findings of the museum's director, Tomasz Kranz (2005), concerning the actual number of camp victims. It would also be a good example of "unintentional falsehood" in the information presented in the exhibition and the narrative of the guides. I am referring to objective facts being presented to visitors inaccurately, although without the intention of misleading them. This applies to fundamental information about the functioning of the camp and is clearly present in visitors' questions (Kutnik, 2019; Kucia, 2005).

It should be stressed that from the start of research on the issue of the number of victims, there has been a major problem with developing an appropriate methodology to estimate this value (Kranz, 2005). The first estimates, based on the findings of the Polish-Soviet Commission, were characterized by quite

a significant discrepancy and spoke of a range between two and one and a half million deceased (the information in the first exhibition spoke of about two million deceased). Since the camp documentation, including the records, was largely damaged by the Germans, a different way to find out the truth about the numbers was needed. One of the methods adopted was based on counting the pairs of shoes left in the camp grounds (which suggested about one million victims). The first revision, based on an analysis of available but scarce evidence material, was made in 1948 and the number of deceased was established as 360,000 (including 200,000 Jews). Subsequent revisions, based on the confrontation of various sources, took place in 1992 (235,000 people dead, including 120,000 Jews) and in 2005. At that time it was established and now is accepted (obviously, this is still not a definite and absolute figure) that 78,000 people died at Majdanek (59,000 of them Jews). The relevant values are not so much true as probable.

3. Authenticity of the place

When considering the issues presented here, a certain distinction should be made, namely between authenticity in the context of the place and space, and authenticity in relation to the truth conveyed in the arranged exposition and the narrative presented there. In both cases, I will try to show the problems by using only a few selected aspects as examples. The former theme seems to be interesting, as the area of the State Museum at Majdanek consists of only one-third of the area originally designated for the camp managed by the Germans (Olesiuk, 2011). It is worth seeing how the various concepts of its arrangement have evolved. First of all, one should consider what is meant by the term "the real space of Majdanek."

KL Lublin was liquidated starting from April 1944, which involved transporting the remaining prisoners and destroying the documentation (Olesiuk, 2011; Kranz, 2005). At the time, dismantling of the facilities and systematic blurring of traces began. However, the front was approaching – one can assume – surprisingly quickly for the German occupants. The last group of prisoners was evacuated as late as July 20. And during the day on July 22, executions took place on the camp grounds. Majdanek was liberated by the Red Army already during the night of July

22–23 (the German garrison in Lublin capitulated a day later). Within the Majdanek compound were unburied and uncremated bodies of victims and a group of living war-disabled Russians whom the SS guards had not managed to evacuate or kill. The first visitors appeared on July 23. The liberated camp was immediately taken over by the Red Army, who set up a recruitment point for the refugee population, a place of isolation for Volksdeutsche, collaborators and Germans, as well as for soldiers of the Polish Home Army (Armia Krajowa) and Farmers' Battalions (Bataliony Chłopskie). Local decision-makers soon noticed the need to protect the site from the threat of rapid devastation. As a result, the institution was designated as a museum and actually started to operate already in autumn 1944 (Regulation on handing over the land and facilities of the former camp to the Polish-Soviet Commission in order to create the Museum, in: Rajca, Wiśniewska, 1984, pp. 9–10).

As mentioned earlier, when speaking about "the truth of the place" in relation to KL Lublin, it is worth considering more carefully what we mean by this notion. Undoubtedly the spatial arrangement, buildings and all installations erected from the beginning of the camp's functioning (from autumn 1941) and existing until July 22, 1944 can be described as authentic or real (Olesiuk, 2011). The problem arises as to what should be done about acts of destruction of buildings carried out by the Germans themselves. How do we assess them? The retreating SS brought about, for example, the burning down of the crematorium building, and if we want to express their intentions, pragmatic barbarity and above all the desire to obliterate traces, it is worth considering whether just the skeleton of the building should have been left and displayed (whereas in fact the crematorium was one of the first buildings to be reconstructed).

Leaving nuances aside, rearrangement of the space (which in itself has important value as historical testimony) already began from July 23. Majdanek – again playing a pioneering role – became the first German concentration camp used by one of the Allies as a prison, intended for the recent abusers, prisoners of war but also soldiers of the Polish underground. The Red Army initiated the process of deconstruction or decomposition of the camp infrastructure (in particular by selling wood from the walls of the barracks to local farmers). Provision shortages pushed soldiers to plunder everything with any utility value. Therefore,

camp property was massively devastated during the first weeks of freedom. Moreover, to use a metaphor, in the newly "used" part of Majdanek, "overwriting" of the place started (like painting a new picture on an existing painting treated as a canvas). Inscriptions carved by prisoners of the NKVD camp began to cover previous ones, carved into the same boards, but by the victims of the Nazi camp. Both of them, from the moment they were made, became full-fledged testimony of terror.

Listing other fundamental problems related to the issue of "the authenticity of the place" and taking into account what visitors are especially interested in (Kutnik, 2016–2019), first of all it is necessary to point to the question of the camp's location in the urban space. During the Nazi occupation it was located on the outskirts of the city, although at the same time along one of the main arteries leading from the center of Lublin to the east. However, today we can say the undeveloped area of the museum is even one of the "dominants" within the city limits, due to the huge area excluded from regular urban development. The current location is quite misleading. Visitors often ask how Lublin residents could have lived here, since the camp is practically located within the city (Kutnik, 2019). What happens here is a peculiar and completely ahistorical projection of the observed environment, where the museum area is surrounded by residential buildings on three sides. Despite the fact that the camp was originally located in the city, during its functioning the few inhabited farms around it were situated some distance away. However, it should be stressed that KL Lublin was by no means conceived by the Nazis as a hidden camp, deliberately located in an inaccessible area. City residents were aware of the horror of what was going on behind the barbed wire fence.

Another important issue related to the postwar shaping of the place (Olesiuk, 2011) was the original and changing arrangement plans (first afforestation and then deforestation of the land, necessary to save the camp relics). Nowadays, probably the most important issue is the conservation and reconstruction of objects and facilities which were not intended by the designers and builders of the camp to be durable. For many years the museum, euphemistically speaking, waged an uneven war against unfavorable circumstances, the impermanence of matter and the passage of time. After the initial attendance success of the museum's first three years, at the turn of the 1940s and 1950s there

was a clear decline in the number of visitors. Auschwitz-Birkenau dominated both mass awareness as a symbol of Nazi genocide and the will of decision-makers to promote one institution at the expense of others (Cebulski, 2016). At that time ideas appeared for a different arrangement of the space of Majdanek, including its liquidation (Olesiuk, 2011) and the construction of a residential district there (for comparison, similar plans were made in Sztutowo, i.e. to set up a vacation center on the Stutthof camp grounds, see Wóycicka, 2009).

A vicious circle appeared, threatening the museum's very existence and the preservation of the site of memory. The consequence of the depreciation of the museum's importance was its long-term underfunding, which resulted in a progressive decrease in the number of visitors. In order to save itself from a drastic lack of funds, the museum had to commercialize part of its land by renting some of the camp barracks to, among others, Motozbyt (an automotive company) and Lubelskie Przedsiębiorstwo Obrotu Zwierzętami Rzeźnymi (a Lublin-based company trading in slaughter animals). Attempts were made to use the camp space, although not the part where prisoners had been held, as pragmatically as possible. In fact, the museum adapted "several barracks to serve as a canteen and a small hotel for visitors" (Olesiuk, 2011, p. 243)! At the same time, there was also a small housing estate in the former SS barracks, intended mainly for employees and their families.

4. Truth and falsity of the exposition – "poisoned pen" of propaganda

The most interesting and probably the most telling theme among the issues discussed here is the fully intentional falsification of "historical truth" for propaganda purposes (Wóycicka, 2009; Banach, 2013, 2014, 2016). Obviously, the content and even the form of the exposition always correspond in some way to the conditions of current history-based politics, and conscious curators deal with this influence with better or worse results. However, the first years of the Museum at Majdanek were a period of particularly strong and not very sophisticated manipulation. I will treat the first exhibitions, from the years 1944–1950, as a symptomatic example of this phenomenon. On one hand, they

were a completely pioneering answer to the question of how to depict the enormity of Nazi genocide in a museum space. On the other hand, undoubtedly they were also a cog in a huge propaganda machine. Their authors had to make compromises on many levels, at the expense of truth, meeting the expectations of the authorities in order to make effective use of the propaganda potential of the place and the history behind it.

Anna Ziębińska-Witek (2011) analyzes the first exhibition in the theoretical context of the "representation of real death," at the same time showing how far it was from accepted museological standards expressed in the concepts of Volkhard Knigge (2003) or Bohdan Rymaszewski (2000). In their view, the role of the exhibition is subordinate in relation to the place, and at the same time new or added elements should be limited to the necessary minimum. Both authors also believe it is practically inadmissible to reconstruct the museum fabric, as well as to reconstruct events. The essence of this approach seems to be perfectly expressed by Rymaszewski (ibid., p. 89–90), who declares that:

> Any ideas connected with constructing a realistic picture of past events on camp grounds should be rejected categorically. Proposals suggesting e.g. setting up wax figures in order to present the participants of the camp drama in a more suggestive way are wrong in their assumption. Putting them into practice would be a violation of the seriousness of the place, its profanation and disrespect for the memory of the victims.

Referring to the above thought, Ziębińska-Witek (2011, p. 162) accuses this approach of being, in spite of its "sincere and noble" intentions, simply too idealistic in that it assumes that authenticity and the aura of the place in itself is enough to have an effective impact on visitors (see Ziębińska-Witek, 2015).

The first exhibition at Majdanek violated and even broke the basic principles expressed in the above-mentioned concept, and it violated them in a situation when the space itself was definitely "more authentic" (a large number of preserved facilities and objects). Access to witnesses was very easy, and the ashes of the victims were still being collected from various places in and around the camp. In these circumstances, nothing more seems to be needed, and the situation appears ideal from a curator's viewpoint. However, in the case of that first exhibition, the means of expression used give the impression of

being very inadequate or even inappropriate (Ziębińska-Witek, 2011; Banach, 2014; Wóycicka, 2009). The exhibition's design was overwhelming and its creators used mainly an aesthetic of the macabre and horror. This form of display was quite quickly found to be unacceptable, although, as Ziębińska-Witek (2011, p. 164) emphasizes, of course it cannot be judged ahistorically and the choice of such a composition finds some justification (or rather understanding) in the postwar period. Certainly, the public at that time was still insensitive to drastic performances. It was also a time when the exhibition at Majdanek was not treated as "historical" but above all as "evidence in the case" against the Nazis (Ferski, 1944/1984). It constituted proof which simply seemed more suggestive in such a form.

The history of the exhibition at the State Museum at Majdanek has been thoroughly analyzed by Krzysztof Banach (2014). In his extensive study, he writes that the first exhibition was opened on Sept. 2, 1945, and only six photographs and a description of the concept have survived to the present day. Thus, we know that the exhibition was dominated by two elements: mannequins dressed in prisoner clothes, with real, individual numbers of victims, and cans with Zyklon B on which these figures were placed. In addition to this, a real human skull was displayed on one of the shelves. We also know that the staging, though it was realized only in part, was carefully planned and no element in it was accidental. This is indicated by the peculiar description of the exhibition concept prepared by Antoni Ferski (the museum's first director), who clearly expressed both the current state of knowledge and "the demand" the exhibition was supposed to meet (Ferski, 1944/1984, p. 12):

> What happened in the death camp will be made expressive, as if coming to life, by setting up a whole series of wax figures at all the sites of torture.
> I am designing the arrangement of the wax figures in scenes, such as:
> 1. Several dozen people dressed in striped clothes are standing in front of a Gestapo officer holding a whip, and next to him a dog is tearing apart the body of one of the unfortunates.
> 2. A Gestapo officer pulling a newborn from its mother's arms.
> 3. Several dozen people are lying in a gas chamber, poisoned by "Zyklon," and many more similar scenes.
> Such scenes from the life of the camp, reconstructed with the use of wax figures, will illustrate the enormity of the horror that prevailed

in the camp at Majdanek. Viewers visiting the Museum will experience the horror of a day in a camp that was set up to exterminate "non-Germans."
The Museum at Majdanek and the evidence gathered there will be a weighty accusation against the pack of murderers and exterminators of humanity, German fascists, before the world.

I think it is difficult not to notice that such a form of representation came dangerously close to the unintended effect of grotesqueness.

In the following years the museum management lost its autonomy rather than showing initiative. The narrative was imposed from above. This is clearly visible in another symptomatic example of a "creative" approach to historical truth that is strongly reminiscent of the actions of the Ministry of Truth from George Orwell's *1984*. Namely, the state of Polish-Jewish relations at a given time significantly influenced changes in the content of the exhibition (Banach, 2013, 2016). Quite tense anyway, in the case of museums at Shoah sites these relations manifested themselves in the intensification of a "rivalry of memory" (Wóycicka, 2009). Already in the first exposition, the dominant message was that the largest number of those killed were Poles, and information about the Jewish nationality was practically absent from the discourse relatively early (Banach, 2014). As a result, the narrative was that there were no Jews at Majdanek and it is primarily a space of Polish and Soviet (or Soviet and Polish) victims. And while it is true that, listing them according to group size, the victims of Majdanek were Jews, Poles and Belarusians, by "splitting" the first group according to prewar citizenship we arrive at a completely new estimate – a peculiar historical half-truth. The issue of nationality was presented in this way at the central point of the exhibition, i.e. on a map showing the 28 nations (with "unassigned" Jews as a separate group) represented among the murdered prisoners of Majdanek. The map was also the focal point of a presentation about KL Lublin which toured Poland.

On the other hand, in the following year (1946), when the museum management tried to emphasize the internationalization of victims in an exposition in which a space for particular national groups was assigned in Prisoner Field III, in the barracks located there, the only realized project was the one devoted to Jews (prepared by the Central Committee of Polish Jews). It was on

display for a relatively short period, probably until 1949. Banach (ibid., pp. 281–282) describes this project as follows:

> The exhibition in the Jewish barrack was presented in a solemn convention similar to the permanent exhibition from 1945. However, it was much more modest in terms of resources. The main idea was to create a mood of reverie and mourning. The photographs and statistics shown in the barrack were only an introduction leading visitors to a symbolic catafalque or sacrificial altar on which, apart from a standing menorah, inscriptions and flowers, drapery and scarfs, a picture of Rachel mourning the Jewish nation was painted.

Another clear example of propaganda falsification of true knowledge about Majdanek was found in the description of Operation Reinhardt, the plan to completely exterminate Jews, and only Jews, in the General Government and plunder their property (Banach, 2013, 2014, 2016). The operation's headquarters was located in Lublin, and Majdanek was where objects belonging to the murdered were gathered. In the description of the exhibition devoted to these events, instead of the word "Jew" the terms "prisoners," "victims" or "people" were used. This universalization of the crime obliterated its essence, i.e. being a victim only because one was a Jew. As for the use of the term "extermination," it was given a new connotation: "the extermination of the nations of Europe," within the framework of which the plan to exterminate Jews was only an initial stage.

The year 1950 brought significant changes in government "martyrdom policy," which resulted in the progressive falsification of the "truth of the place" in relation to Majdanek (Banach, 2014). The new remembrance policy is reflected in the instructions of the Central Committee of the Polish United Workers' Party (the ruling communist party) for the Polish Association of Former Political Prisoners of Nazi Prisons and Concentration Camps, which stated: "Commemoration of the victims of Nazi terror should be kept within healthy and necessary boundaries and should focus on combat aspects, not on suffering." The document also stressed that "political prisoners should not be treated as priests of *martyrdom* but as conscious and active members of society" (ibid., p. 283). This tone leaves little room for any far-reaching interpretation of the guidelines. In the 1950s the national identification of victims was completely omitted from the narrative used in exhibition descriptions.

A change of focus to the need for commemoration, martyrdom and its social function in the following years even threatened to cause the museum to be closed down. The truth of the place ceased to be any bargaining chip in this respect. On the other hand, the unquestionable falsification of the exposition's message made it practically worthless, both in the context of expanding knowledge about the Holocaust and Nazi terror in general and as a form of commemoration of the victims. After the Khrushchev Thaw (1956), according to Banach (ibid., p. 273), the museum presented "as much reliable knowledge as the situation allowed." But it is also true that a period of such strong, top-down control of the content did not happen again.

5. "A museum more real"

I would like to conclude the "case study" proposed here by moving on to the current problems signaled at the beginning, which are related to the question of preserving the "authenticity of the site," while also invoking the earlier call expressed by a visitor for the museum to be "more real (authentic, true)." Significant and relatively recent changes in the arrangement of the exposition were caused by inappropriate and irresponsible "actions" of visitors themselves. Major incidents from recent years which forced the museum management to intervene and modify the form of representation include, for example, a smiling little girl photographed inside a crematorium furnace (KAD, 2008), a picture by Carl Michael von Hausswolff which he claimed to have painted with ashes collected within the Majdanek compound (Franczak, 2012), and the theft of shoes presented in the exhibition, belonging to victims (Szlachetka, 2014).

The effect of this shocking behavior or even profanation of the site by visitors was the progressive "closing" of the exposition using the simplest and at the same time most effective means. Access to prisoners' shoes and bunks was restricted by barriers while the crematorium furnaces were fenced off with glass. The ultimate effect is that the space has become "unrealistic" in a sense. A paradoxical "museumization" or "gallerization" of this particular museum (where the most important means of expression seems to be the authenticity of both the space and the objects on display) has taken place. This is also reflected in comments from people visiting Majdanek who remember

(and mention) how the space directly available to them during their earlier visits affected them (Kutnik, 2016–2019). For example, it made a great impression on people when they could walk between piles of shoes or bunks (experiencing the huge barrack and the distinctive lack of breathing air in a building with no windows), or when there was no barrier between them and the crematorium furnace and the whole murderous installation was described as more tangible.

Can the remedy be a return to greater authenticity? How would this be achieved when the last witnesses are passing away and the museum space needs to be protected from the visitors in the first place? For years now, human remains have not been used in the exhibition. It is hard not to agree with those who treat the use of human remains as unethical and permanent objectification of the victims (Ziębińska-Witek, 2011). Of course, no new material substance will appear in the form of buildings or camp relics. On the contrary, there will be fewer and fewer objects. The camp kitchen barrack located in Prisoner Field III self-ignited and burned down in 2010 (Olesiuk, 2011). There was a discussion on how to deal with the damaged facility. The museum management opted for leaving this space in the form that was left after the fire. The burnt building was itself an early reconstruction. Out of five prisoner fields, not one is complete or fully reconstructed. Asking specifically how much of KL Lublin remains in the museum space and exhibition at Majdanek, based on recent research the answer will be that it is about 25% of the originally existing facilities on one-third of the original area (which totaled 270 hectares). On the other hand, of the presently existing "physical" facilities, around 40% is a reconstruction.

Skillfully using the authenticity of the place and finding new forms of "speaking" to visitors is therefore a great challenge facing exhibitors. The value of Majdanek, Auschwitz-Birkenau, Stutthof and other museums in World War II camps is inestimable and indisputable. If Emmanuel Lévinas was right (2012) that the noble slogan "no more war" can be achieved only by a fundamental change in mindset, by any means possible, but mainly through education using a rich list of various cultural means, then the potential of sites of memory, authentic facilities, objects and spaces is enormous and still leaves plenty of room for action. After all, we should keep in mind at all times

the alarming memento that "the end of history" has not even started yet. Nowadays, to make the ethics of responsibility for the Other a truly dominant "first philosophy," in the face of the memory of the boundless cruelty of the Holocaust, in an educational aspect, these particular institutions play a key role and are irreplaceable.

References

Banach, K. (2014). Działalność wystawiennicza Państwowego Muzeum na Majdanku w latach 1944–2014 [Exhibition work of the State Museum at Majdanek in 1944–2014]. *Zeszyty Majdanka, XXVI*, 273–310.

Banach, K. (2016). Exhibiting violence or teaching values? Historical exhibitions at a modern museum of martyrdom. *Przegląd Historyczny, 107 (1)*, 77–1.

Banach, K. (2013). Zagłada Żydów w świetle ekspozycji stałych Państwowego Muzeum na Majdanku [Presentation of the Holocaust at permanent exhibitions of the State Museum at Majdanek]. *Studia Judaica*, 16 (2), 115–143.

Cebulski, T. (2016). *Auschwitz po Auschwitz. Polityczne i międzynarodowe aspekty funkcjonowania Państwowego Muzeum Auschwitz-Birkenau 1980–2010* [Auschwitz after Auschwitz. Political and International Aspects of the Functioning of the Auschwitz-Birkenau State Museum 1980–2010]. Kraków: Wydawnictwo LIBRON.

Ferski, A. (1944/1984). Raport o organizacji Muzeum [Report on the organization of the Museum]. In: C. Rajca, A. Wiśniewska (Eds.), *Powstanie Państwowego Muzeum na Majdanku. Wybór dokumentów* [Establishment of the State Museum at Majdanek. Selection of Documents] (pp. 12–13). Lublin: Państwowe Muzeum na Majdanku.

Franczak, P. (2012). Wszczęto śledztwa w sprawie obrazu namalowanego prochami ofiar Majdanka [Investigation initiated on a painting painted with the ashes of Majdanek's victims]. *Kurier Lubelski*, Dec. 10.

KAD (2008). Zdjęcie z Majdanka do prokuratury [A photo from Majdanek reported to the prosecutor's office]. *Gazeta Wyborcza Lublin*, Sept. 13.

Knigge, V. (2003). Teren byłego obozu a wystawa historyczna [The area of the former camp and the historical exhibition]. In: K. Marszałek (Ed.), *Chronić dla przyszłości. Międzynarodowa Konferencja Konserwatorska, Oświęcim, 23–25 czerwca 2003 roku* [Protect for the Future. International Conservation Conference, Oświęcim, June 23–25, 2003] (pp. 114–118). Oświęcim: Państwowe Muzeum Auschwitz-Birkenau.

Kranz, T. (2005). Ewidencja zgonów i śmiertelność więźniów KL Lublin [Records of deaths and mortality of KL Lublin prisoners]. *Zeszyty Majdanka, XXIII*, 7–54.

Kranz, T. (2003). Muzea w byłych obozach w Polsce jako forma instytucjonalizacji pamięci [Museums in former camps in Poland as a form of institutionalization of memory]. *Łambinowicki Rocznik Muzealny, 26*, 9–22.

Kranz, T. (2011). O koncepcji modernizacji ekspozycji Państwowego Muzeum na Majdanku [On the concept of modernization of the exhibition at the State Museum at Majdanek]. *Zeszyty Majdanka, XXV*, 9–23.

Kucia, M. (2005). *Auschwitz jako fakt społeczny. Historia, współczesność i świadomość społeczna KL Auschwitz w Polsce* [Auschwitz as a Social Fact: The History, Present, and Social Consciousness of KL Auschwitz in Poland]. Kraków: Universitas.

Kutnik, J. (2016–2019). Unpublished materials from the hospitations of groups guided around the State Museum at Majdanek; research carried out in 2016–2019.

Lévinas, E. (2012). *Całość i nieskończoność. Esej o zewnętrzności* (Polish edition of *Totalité et infini. Essai sur l'extériorité*, transl. M. Kowalska). Warszawa: Wydawnictwo Naukowe PWN.

Moore, K. (2000). *Museums and Popular Culture*. London: Cassell.

Nowacki, M. (2015). Autentyczność w percepcji osób zwiedzających muzeum [Authenticity in the perception of museum visitors]. In: J. Hochleitner, W. Połom-Jakubowicz (Eds.), *Funkcje muzeum współcześnie. Wyzwania i interpretacja oferty muzealnej* [Functions of the Museum Today. Challenges and Interpretation of the Museum Offer] (pp. 91–103). Malbork: Muzeum Zamkowe w Malborku.

Olesiuk, D. (2011). Urządzenie przestrzenne Państwowego Muzeum na Majdanku 1944–2010. Plany i realizacja [Arrangement of the space of the State Museum at Majdanek 1944–2010. Plans and implementation]. *Zeszyty Majdanka, XXV*, 227–271.

Rymaszewski, B. (2003). Granice ingerencji muzealnej i konserwatorskiej w Muzeum i Miejscu Pamięci Auschwitz [Limits of museum and conservation interference in the Auschwitz Memorial and Museum]. In: K. Marszałek (Ed.), *Chronić dla przyszłości. Międzynarodowa Konferencja Konserwatorska, Oświęcim, 23–25 czerwca 2003 roku* [Protect for the Future. International Conservation Conference, Oświęcim, June 23–25, 2003] (pp. 24–34). Oświęcim: Państwowe Muzeum Auschwitz-Birkenau.

Rymaszewski, B. (2000). *Pamiętać będą pokolenia* [Generations Will Remember]. Oświęcim: Państwowe Muzeum Auschwitz-Birkenau w Oświęcimiu.

Szlachetka, M. (2014). Państwowe Muzeum na Majdanku: Kradzież butów więźniarskich [The State Museum at Majdanek: The theft of shoes belonging to prisoners]. *Kurier Lubelski*, Nov. 25.

Szlachetka, M. (2015). Zdjęcie byłego obozu koncentracyjnego na Majdanku jako fototapeta [A photo of the former Majdanek concentration camp as photo wallpaper]. *Kurier Lubelski*, Aug. 25.

Unger, P. (1988). *Muzea w nauczaniu historii* [Museums in History Teaching]. Warszawa: Wydawnictwa Szkolne i Pedagogiczne.

Webber, J. (2003). Znaczenie fizycznych śladów przeszłości dla edukacji współczesnego społeczeństwa [The importance of material traces of the past for the education of modern society]. In: K. Marszałek (Ed.), *Chronić dla przyszłości. Międzynarodowa Konferencja Konserwatorska, Oświęcim, 23–25 czerwca 2003 roku* [Protect for the Future. International Conservation

Conference, Oświęcim, June 23–25, 2003] (pp. 104–113). Oświęcim: Państwowe Muzeum Auschwitz-Birkenau.

Wóycicka, Z. (2009). *Przerwana żałoba. Polskie spory wokół pamięci nazistowskich obozów koncentracyjnych i zagłady 1944–1950* [Arrested Mourning: Memory of the Nazi Camps in Poland, 1944–1950]. Warszawa: Wydawnictwo TRIO.

Zarządzenie o przekazaniu terenu i obiektów byłego obozu Polsko-Radzieckiej Komisji w celu utworzenia Muzeum [Regulation on handing over the land and facilities of the former camp to the Polish-Soviet Commission in order to create the Museum]. In: C. Rajca, A. Wiśniewska (Eds.), *Powstanie Państwowego Muzeum na Majdanku. Wybór dokumentów* [Establishment of the State Museum at Majdanek. Selection of Documents] (pp. 9–10). Lublin: Państwowe Muzeum na Majdanku.

Ziębińska-Witek, A. (2011). *Historia w muzeach. Studium ekspozycji Holokaustu* [History in Museums: A Study on Holocaust Exhibitions]. Lublin: Wydawnictwo Uniwersytetu Marii Curie-Skłodowskiej.

Ziębińska-Witek, A. (2015). Muzea wobec nowych trendów w humanistyce. Refleksje teoretyczne [Museums in the face of new trends in the humanities: Theoretical reflections]. *Historyka. Studia Metodologiczne, 45*, 97–115.

Ziębińska-Witek, A. (2009). Problemy wizualizacji Holokaustu – ekspozycje w muzeach miejscach pamięci [Problems of Holocaust visualization – exhibitions in museums and memorial sites]. In: P. Chruścielski, M. Owsiński (Eds.), *Konferencja „Od Westerplatte do Norymbergi. II wojna światowa we współczesnej historiografii, muzealnictwie i edukacji"* [Conference: From Westerplatte to Nuremberg. World War II in Contemporary Historiography, Museology and Education] (pp. 293–296). Sztutowo: Muzeum Stutthof.

Index of Names

The authors discuss truth and falsehood in science and the arts. They view truth as an irreducible point of reference, both in striving for elementary knowledge about the world and in seeking methods and artistic means of achieving this goal. The multilevel and multiple-aspect research presented here, conducted on material from different periods and different cultures, shows very clearly that truth and falsehood lie at the foundation of all human motivation, choices, decisions, and behaviors. At the same time, however, it reveals that every bid to extrapolate the results of detailed studies into generalizations aimed at universalization – by the very fact of their discursivation – either subjects the discussion to the rules of formal logic or situates it outside the realm of truth and falsehood.

The Editors