ACTA PHILOSOPHICA FENNICA

Vol. 99

Past. Future. Philosophy.

Proceedings of the Philosophical Society of Finland Colloquium 2022

Edited by
EERO KAILA,
HENRI PETTERSSON
& JANI SINOKKI

HELSINKI 2023

ISBN 978−951−9264−98−1
ISSN 0355-1792

Hansaprint Oy
Helsinki 2023

Simo Knuuttila 1946–2022

Table of Contents

Introduction

The immediate cause for the existence of this collection stems from the annual colloquium of the Philosophical Society of Finland, organized at the University of Oulu in June of 2022. However, the results displayed here should not be regarded as straightforward conference proceedings, even if the essays are indeed based on presentations premiered at that event. First, what we find here is only a small proportion of the total of 29 presentations seen at the colloquium. Second, the following papers are not merely the written transcriptions of the talks, as the authors have expanded their presentations to the length of full research articles. As such, the manuscripts submitted for publication went through a further round of blind peer review during the editorial process. We thank everyone who submitted their paper for consideration and the anonymous peer reviewers for their constructive assessments.

The annual colloquiums of The Philosophical Society of Finland are organized as international researcher seminars, with free theming, on alternate years. Every other year, the colloquiums are aimed for the Finnish speaking general audiences. This colloquium 2022 was the third international event so far, with the previous conferences being the Philosophical Society of Finland Colloquium 2018 (PSFC 2018) and Annual Colloquium of the Philosophical Society of Finland 2020 (FiPhi 2020). The overarching theme for the 2022 meeting was given in the title "Past. Future. Philosophy." The idea was to highlight the strengths and the two focus areas of the University of Oulu as the host organization, namely the philosophy of history and the philosophy of education—in some sense, the former looks to the past and the latter to the future. However, this loose theme allows for a wide range of philosophical topics and approaches, as the results here show. A workshop titled "The Future of Philosophy of History" was

held concurrently with the colloquium, organized by Jouni-Matti Kuukkanen, Georg Gangl and Ilkka Lähteenmäki. Various presentations from the workshop have been published in a special issue "The Futures of Philosophy of History" of *Journal of the Philosophy of History* (Volume 17, issue 2, 2023), edited by Gangl and Lähteenmäki.

The passing of Professor Simo Knuuttila only a week after the colloquium casts a sad shadow over our memories of the colloquium. Knuuttila was the guest of honor at the event, delivering the opening keynote presentation. To honor his memory, his presentation is included here. In his paper, Knuuttila reflects on the expectations various philosophers have placed on the field of the history of philosophy. In his text, he presents and comments on a wide range of approaches to understanding the nature and purpose of history of philosophy.

The second keynote presentation is Miira Tuominen's "Why the Past is Important for the Future in Philosophy". In it Tuominen reflects on the different considerations we must pay attention to when we approach historical philosophers. In conclusion she provides a defending argument in case for availability of the Classics against unnecessary censorship based on contemporary, political anxieties: learning from past authors, flawed as they may be, is essential for our understanding of the present.

In the opening article of this volume, Eerik Lagerspetz and Sari Roman-Lagerspetz offer a detailed look at the much discussed, and much criticized, thesis of "end of history". The thesis was popularized by Francis Fukuyama, but as Lagerspetz and Roman-Lagerspetz show, it originates in works of G. W. F. Hegel, and especially in the readings of Hegel by Russian-French philosopher Alexandre Kojève. Without taking a definite stand on right way to read Hegel, the authors argue that while many authors before Kojève discussed the idea of end of history, it is Kojève how first thought that Hegel's philosophy actually implies this thesis, and furthermore, Kojève is the first philosopher to argue that the thesis also is true. As Kojève's interpretation of this thesis as well as Hegel has been very influential, the authors here provide a valuable critical discussion about the historical details of this influential philosophical thesis.

In his article, Sami Pihlström draws on his recent book developing a pragmatist philosophy of the humanities. Pihlström is interested not in methodological questions of historiography, but more specifically in the general realism issue as applied to the study of the past from a pragmatist perspective. His focus is pragmatist view of truth, and its application to studying the historical past. Pihlström points out several specific, but important issues and concepts central to the pragmatist approach in historiography, especially relating to the important connection between epistemology and ontology, which raises many questions not only for pragmatist approaches, but for philosophy of history more generally.

Olli Lagerspetz discusses Peter Winch's view in relation to the preconditions of making truth claims about the past. The issue Lagerspetz focuses is specifically on against what kind of general background any meaningful knowledge claims about the past could be made, for in general, such claims presuppose at least a certain systematicity to the working of nature. Focusing on Winch's interests in metaphysical questions about truths about the past, yet inside the framework of Wittgenstein's philosophy that is often considered hostile to such metaphysical questions, Lagerspetz questions some important standard realist presuppositions about knowledge of the past and its relation to the knowledge about the present.

Paavo Pylkkänen offers the reader a detailed exposition of physicist David Bohm's "implicate order" world view. That view implies a physics-based way to think about time and movement that is interesting to philosophers as well in their own right. Pylkkänen shows the way Bohm's implicate order view challenges the traditional "block-universe" view and replaces it with a view of past as "enfolded structures" and future as "unfolded structures" in the present moment. Pylkkänen indicates how Bohm's view has implications for physical theories, such as avoiding postulating particles moving backwards in time, but also philosophically interesting consequences as well, for instance, about the nature of experiences about movement and passage of time.

In his chapter, Jani Hakkarainen analyzes the relationship between metaphysical realism and realism about metaphysical realism. He argues, by drawing on the formal ontological tradition and by looking at several thinkers in that tradition,

that the two views should be distinguished and kept apart. Hakkarainen discusses the relation between the views and emphasizes how arguments against standard metaphysical realism (e.g., such as were presented by David Hume) cannot be settled on their own without entering argumentation about the second-order view about metaphysical realism about metaphysics as well. As Hakkarainen proceeds, he suggests theoretical ways to relate different areas of metaphysical study to each other to present a general topology of this complex field of study.

Henriikka Hannula provides a systematic and step-by-step exegetical reconstruction of Dilthey's theory of historical understanding, which offers an answer to the question of how we can know anything about the past. Hannula first contextualizes Dilthey's project in its proper intellectual-historical setting and then reviews the main epistemological and ontological arguments Dilthey gave for the possibility of scientific historical understanding. As a result, we acquire a comprehensive picture of Dilthey's philosophical thinking on the conditions and possibilities of historical knowledge.

In "A Note to 'Meaning in Time'" Jaakko Reinikainen examines "Kripkenstein's" skeptical challenge in the field of philosophy of language. Reinikainen considers how the causal-historical account of reference put forward by Kripke and Devitt can help us solve the challenge in two steps. First is to see the challenge not as a metaphysical but as a logical problem that leads us to accept a temporal, historical account of meaning. Second is to see meaning as an aspect of causal-historical reference, in which case meanings will be understood as historical in nature.

Katariina Lipsanen analyzes Henri Bergson's views regarding the temporal nature of philosophy. In this undertaking, Bergson's concept of duration (*durée*) is of central importance, as it can help us see how Bergson's philosophical methodology carries far-reaching ramifications for his overall conception of philosophy. Lipsanen also provides examples of how these ideas might have practical applications.

Johan Sandelin discusses various ways to understand and to tackle the Non-Identity Problem. He also asks, whether we have a moral obligation not to act in a way that triggers injustices? As a solution, Sandelin looks at cases involving seem-

ingly unalterable premises. He advances a justice-based principle that by considering unjust also those fates that could not be bettered, provides a novel approach to the Non-Identity Problem.

In his paper "The future of political institutions" Tuukka Brunila reconsiders the normative dimensions of Michel Foucault's genealogical method, especially as they pertain to analyzing the nature of political institutions. He puts forward a normative reading of Foucauldian genealogical critique, and through its lens views political institutions as historically contingent phenomena. This helps us to see how political institutions are transformable.

In addition to the editors of this volume, the organizing committee for the colloquium consisted of Professors Jouni-Matti Kuukkanen, Katariina Holma and lecturer Vili Lähteenmäki. The organizers would also like to thank Veli-Mikko Kauppi, Minna-Kerttu Kekki, Anna Itkonen and Samuel Iinatti for their assistance with the practicalities in running the event.

Eero Kaila, Henri Pettersson & Jani Sinokki

Keynote Presentations

The History of Philosophy – What can be expected from it

SIMO KNUUTTILA[1]

Western philosophy has not been especially time-conscious in its acts of self-reflection. This goes without the obvious references to what has previously been said about the subject. Philosophy, in itself, is neither the practice of expecting and anticipating the future, nor of remembering the past. Works of philosophy contain defended claims which are assumed to hold ground in perpetuity, as Kierkegaard says when talking about existential choices. When temporality and its categories are discussed within philosophy, the viewpoint is thought to be external to time. This is the case with Kant's transcendental philosophy as well as Heidegger's existential analysis in his book *Being and Time*. Perhaps one could ask, how could it be otherwise? Even though practitioners of philosophy are temporal and changing beings, their cognitive culture is an intellectual activity which considers understandable reality and its conditions as residing in the conceptual present. I shall not address here any committed political philosophy such as messianism or Leninism which would serve to advance a desired future.

A great part of philosophical literature is source material for the history of philosophy. It is, in other words, a record of philosophical works written in past times. Philosophy of an-

[1] Simo Knuuttila held this opening keynote presentation of the Philosophical Society of Finland's Colloquium *Past. Future. Philosophy* in Oulu 9th June 2022. He passed away suddenly only a week after the event. The text is translated into English by the editors. The editors would like to express their gratitude to Marja-Liisa Kakkuri-Knuuttila for making this presentation available.

cient and medieval times has survived in the form of copied manuscripts. Early modern philosophy has persisted also in printed form. The history of philosophy consists of this literature which resides on the shelves of libraries or saved as computer data, even when no one is reading or listening to it in audio book form.

A Western reader well versed in the history of philosophy can usually locate an English translation or written text on the timeline of the history of philosophy, even when the title or the author is unknown to them. What is such identification based on? That the text is part of the history of philosophy might be apparent from its old-fashioned style and the philosophical concepts and possibly some more specific terminology it contains, such as a discussion about general concepts, future contingencies, consequence, determinism, and so on. Naturally, other things are also needed to identify the text. When the text is identified, it can be placed on the right library shelf, if this is all that is required. Many works of the history of philosophy are familiar to practitioners of philosophy in this superficial way.

Deeper and more involved understanding of an older historical volume requires interpretation dictated by temporal distance. There are two traditions here which are called rational reconstruction and historical reconstruction. Rational reconstruction is the most popular mode of interpreting philosophical-historical texts. It was also the approach favored in the early period of so-called analytical philosophy. Russell's book on Leibniz or his later work, *A History of Western Philosophy,* function as examples. In these cases, the text is being read in the light of philosophical inquiry, i.e., it is thought to contain an answer in the form of an argument to a philosophical question in the reader's mind. The task of the interpreter is then to recognize the philosophical problem, to explain how the work answers their question and possibly to evaluate the validity of the answer. Interpretative questions are philosophical if they can be understood in ontological, epistemological, or semantic terms. The relevance of the work in relation to the interpretative questions determines its philosophical nature. According to a common presumption, the subject of interpretation is ultimately philosophical reason. For better or for worse, this same reason is thought to have

directed the original author of the text. The metaphysics of the interpretation then emerges from the general guidelines of philosophy. Philosophical reason seeks rational action from the text. In the tradition of rational reconstruction, investigating the historical background is not actual philosophy but historical research instead.

Historical reconstruction is distinct from rational reconstruction. The former strives toward understanding of meaning. This model was developed in Wilhelm Dilthey's influential philosophy of humanistic sciences, in which the research interest is focussed on historical understanding instead of natural scientific explanation of phenomena. Texts are intended to be read in their original historical context. According to Dilthey, accessing the true meaning of works requires entering the historical horizon of the author. The Western tradition has philosophical connections to the interpreter's own thinking, which makes it comprehensible at first look. To counter, this effective-historical familiarity contains concepts and chains of thought that have another kind of background color, and this requires immersion within the differing as Friedrich Schleiermacher describes. Text can be conceived as the articulation of subjective experience, and the goal of understanding as re-experiencing the original experience, an activity known to enrich the interpreter's consciousness. Historically interpreted truths are understandable in the historical perspective.

The problem of communicating subjective experience was central to discussion spurred by Dilthey's extended methodology, which can be called the hermeneutic theory of interpretation. The key names of this theory in the previous century were Heidegger, Gadamer, and Ricoeur. The question was posed: what would re-experiencing the original mean? The crux of the critical answer was that the original question was already construed from the perspective of the interpreter. The main content of the history of philosophy can be read to be the merging of the text's horizon of meaning and the interpreter's horizon of understanding, taking place within acts of historical interpretation.

A hermeneutic event can be thought of as a kind of hybrid of rational and historical reconstruction because the history of philosophy is defined by effective history and pre-

understanding of tradition. Hermeneutic understanding is described as re-understanding the subject matter of the text. Hermeneutic theory of interpretation has been applied to analytic philosophy by Richard Rorty, who distinguishes its anachronism-averse basic level and, the more applied, Gadamer-type conceptual historical development. Avoiding anachronism is based on Quentin Skinner's philosophical interpretation, which relies on speech-acts. Anachronism, in this context, is defined as a philosophical chain of thought, which is not formulated or intended to be included in the text by the author. How should one think about this? If, for example, Plato's theory of justice is construed, it is understandable that some presupposition of justice works as a starting point. The result of research in this case is such that Plato's concept of justice contains uses of the concept, for example use A and use B. Researcher can also notice that it does not contain an instance of use C, which could be familiar for them in Rawls's theory of justice. To point out this omission is not an anachronism in Plato's case, because on one hand the omission effects only the latter and on the other it shines light on the historical profile of the concept. Alternatively, if the purpose of the researcher was to clarify Plato's concept of justice, multiple references to later ideas are considered inappropriate by many. These have more to do with conceptual history than with Plato's theory. This sort of comparison is not expected from historical research on Plato; on the contrary, many critics consider this strange within the field of Plato scholarship.

The sketch above considers the philosophical-historical interpretation of texts. Since the 1970s, another kind of theory has been present in the debates of philosophy of history (as a scientific discipline): the so-called narrativist theory of historical writing. Its key authority figure, Frank Ankersmit, is primarily interested in political history. According to Ankersmit, essential to prominent works of political history are synthesizing theses that he calls narratives. They shape the internal structure of a presentation and connect it with the conceptual horizon of the interpreter. In Ankersmit's view, the comparison and evaluation of contemporary works of historians is based on politico-moral and esthetic criteria. Epistemic criteria are not a part of evaluative criteria of narrativ-

ist works, because they do not interact with past events as such, nor is truth-seeking description thus an evaluable point of comparison. The narrative structure in presentations concerning the past has to be valued differently, rather than by epistemic interest.

According to Jouni-Matti Kuukkanen, who in his 2015 book *Postnarrativist Historiography of Philosophy* discusses the subject, conceptions of truth are considered to be the weak point of narrativist theories. Kuukkanen writes that evaluation criteria of research include the question of rational justification in relation to its limiting factors. Concerning truth, scholars of history may hold relevant the fact that research is limited by its research material. This requires integrity, comprehensiveness, and coherence. History is distinct from fiction, which, although it also can fulfill the conditions of integrity and coherence, lacks the evidence for argumentation, which is necessary for historical research.

Kuukkanen has also researched the transformation of scientific theories and common presumptions held in conceptual history. He has looked at Lovejoy's history of ideas as well as its critique, and he has defended, in a way, Lovejoy's conception of unit-ideas, although his interpretation has different goals than Lovejoy. Kuukkanen writes that the tradition of a concept can change and remain the tradition of the same concept by assuming on one hand that the same basic element of the concept is retained in between its uses, and on the other that these changes concern matters in its periphery. This model somewhat resembles that found in Aristotle's *Physics,* with its ideas of immutable substances and their changing accidences. While the ideas of Wittgensteinian family resemblance, gradual change, and historical identity are interesting, they are also problematic, as Kuukkanen notes. Should we think that, for example, a concept altered completely in terms of content would be the same as another concept within that tradition? Without defending a metaphysical assumption as strong as this, Kuukkanen holds it reasonable to presume that it is necessary to distinguish the observable core from its mutable part within a conceptual complex. His examples are Koselleck's notes on Bildung and Pauli Kettunen's account of the concept of 'society'. Kuukkanen characterizes his stance on philosophy of history to be that of rational pragmatism.

Jari Kaukua and Vili Lähteenmäki have studied conditions of interpretation. In addition to Kuukkanen's proposed immutable core, Kaukua and Lähteenmäki present in their theory non-textual conceptual explanatory aspects of change. According to their theory, in addition to historical and conceptual reconstruction, the phenomenal level of interpretation needs to be taken into account. Their view is that historical texts contain phenomenal presumptions about the observable world, which constitute the interpreter's everyday experience. Kaukua and Lähteenmäki take note of the phenomenal premiss that is generally accepted in scholarship of history, according to which texts describe (at least in part) the same world that forms the interpreter's world view. According to the 19th century historist theory of interpretation, interpreting a historic event always requires a scientific basis for the interpreted event. This is in accordance with Lähteenmäki's and Kaukua's non-textual aspect of interpretation. If sources report, for example, about strange paranormal miracles page after page, as *the Ecclesiastical History of the English People* by Bede does, they cannot be considered as descriptions of events that have taken place. This, however, is not Kaukua's and Lähteenmäki's example. The historist secular prerequirement is the same as Galilei's principle: physics describes the observable movements of the stars—or whatever they are called.

One of Lähteenmäki and Kaukua's examples is the medieval debate about justified price, in which writers are thought to have a similar pre-conception about the market. One could ask if this consensus is a non-textual presumption, or whether it is in fact a literal assumption that can be read in the popular Aristotelian works. Perhaps the question of proper interest rate would have been more interesting as an example. Why has it been four percent to this date since the medieval times? Lähteenmäki and Kaukua think that temporal-historical discussion in the context of contextual reconstruction is a partly non-textual presumption that can be relevant in terms of changing meaning, even in the case when only parts of the historical texts are known.

Kaukua and Lähteenmäki present an example of the partial overlapping of interpreted text and the interpreter's perspective. This is Avicenna's example where a thirsty person

wants to drink. Interpretations presume in Avicenna's case that this drink is water, which is also extant in the interpreter's world. If the interpreter claims Avicenna meant that the thirsty person wants to drink for example Coca-Cola, one could hold this anachronistic. On the other hand, if it is said that according to Avicenna, a thirsty person wants to drink e.g. Coca-Cola, one could ask why the example is so commercial.

Even though the theories presented above concern the interpretation of the history of philosophy, they are not especially interested in the temporal pre-requirements of interpretation. Actually, this is not even the key point of contention in the scholarship of the history of philosophy. The interpretation of a historical text is presented as contemporary dialog from the viewpoint of its original moment of presentation. The temporal aspect is apparent in the avoidance of general anachronisms. This concerns also the non-conceptual aspects of the text. Some of them, however, are thought to be understandable as speech about the world as it appears to its inhabitants. An interesting detail in this context is the translation of emotion vocabulary. Are the emotional quotes from Cicero or Seneca as such intelligible to viewers of TV shows?

In his book *Leibniz and Arnaud. A Commentary on Their Correspondence* R. C. Sleigh claims that his goal is to illuminate what Leibniz has said, not what he was supposed to say or what he could have said. Sleigh describes analysis to be rational philosophical reconstruction of arguments. Sleigh refers to as rational philosophical reconstruction of arguments the notion of analysis that refers to Leibniz's principles or his displayed habits of thought presented elsewhere, or to contextual principles known in his time. Sleigh's summary is a popular philosophical historical-critical model of research within the Anglo-American tradition. It corresponds to how well-known series and publishers of philosophy conceive the history of philosophy as based on the critical interpretation of historical texts. A work of the history of philosophy can be expected to have philosophically intelligible analysis and to hold historically its place. Narrativist interpretations without the requirement of historical truth are not part of the mainstream in philosophy. The thought that scholarship would

include hermeneutic re-understanding of its subject matter is also not part of the program. This is more akin to non-historical philosophy, which has been inspired by historical sources. Re-interpretation is an essential part of the events of interpretation in Gadamer's hermeneutics. Rorty distinguishes its historical reconstruction as its own part. This sort of act can be a part of interpretation, but the connection is accidental—although not so in Gadamer's reasoning. The lack of hermeneutic development in the history of philosophy is not an issue that would be prevalent in its critique. If the author is not blind to trivial conceptual historical change, it is not presumed that Plato scholarship should figure out how historical ideas should be developed. Historical texts are discussed in the historical horizon as if they were old cars in an authentic historical exhibition. Texts do not require a specific systematic message. Why are they being studied, then? Perhaps general interest toward them is intrinsically valuable, and other interest can derive from personal motivation that authors refer to in the introductions of their books.

Why the Past Is Important for the Future in Philosophy

MIIRA TUOMINEN

1. Spinoza's Excommunication

On July 27 in 1656, Baruch/Benedictus/Bento Spinoza was excommunicated from the united congregation of Portuguese Jews in Amsterdam. Earlier, when still living in Portugal, his family had taken Christian baptism that was made compulsory by the King of Spain. However, they had not abandoned their faith and when they moved to Amsterdam, they joined the Jewish congregation there. Amsterdam was an attractive place at the time. Religious diversity met with international commerce and the city's free atmosphere attracted interesting people. However, it seems that Spinoza's views were too much even there. Although it is plausible that the excommunication was caused by Spinoza's philosophical ideas, some scholars stress that we should not automatically identify the reasons for his excommunication with those views that caused trouble for Spinoza later, especially his famous statement that God is nature that has been taken to represent some kind of atheism.[1]

Incidentally, and much more recently (November-December 2021) when I was preparing for this colloquium intended for January 2022, Rabbi Serfaty from the very same congregation declared a Spinoza scholar, Yitzhak Melamed from the Johns Hopkins university a *persona non grata*. The reason was that Melamed studies Spinoza, a philosopher who was excommunicated because of his views. After that, the synagogue has written back to Melamed and apologized, saying that Rabbi Serfaty did not have the authority to ban him.

[1] Kasher and Biderman (1990, 103).

Rabbi Serfaty agreed to apologize about "the way in which the letter was sent" but sticks to its content.[2]

In the following, I will not discuss Spinoza's excommunication or his philosophy from a scholarly perspective. I am not a Spinoza scholar. Rather, I shall use Spinoza and Aristotle as examples to reflect on the different considerations we need to pay attention to when approaching authors in the history of philosophy. The meaning of this exercise is to show why reflecting on past authors and our attitudes towards them is vital for the future of philosophy.

As we shall see, Spinoza's excommunication was written in harsh words. That was not unusual. However, there are other unusual elements in the affair—for example that the dossier on his case went missing and that we have no primary sources on the discussions that led to the decision.[3] Let us quote some excerpts from the excommunication to get a taste of it.

> (T1) [H]aving long known of the evil opinions and acts of Baruch de Spinoza, they have endeavored by various means and promises, to turn him from his evil ways. But having failed to make him mend his wicked ways, and, on the contrary, daily receiving more and more serious information about the abominable heresies which he practiced and taught and about his monstrous deeds, and having for this numerous trustworthy witnesses who have deposed and born witness to this effect in the presence of the said Espinoza, they became convinced of the truth of this matter; and after all of this has been investigated … they have decided, with their consent, that the said Espinoza should be excommunicated and expelled from the people of Israel…

Spinoza was also thoroughly cursed.

> (T2) By decree of the angels and by the command of the holy men, we excommunicate, expel, curse and damn Baruch de Espinoza, with the consent of God, Blessed be He, and with the consent of the entire holy congregation, and in front of these ho-

---

[2] https://jmoreliving.com/2021/12/03/hopkins-scholar-receives-apology-from-amsterdam-synagogue-that-banned-him/

[3] Kasher and Biderman (1990). I am quoting the excommunication text from their article.

> ly scrolls with the 613 precepts which are written therein; cursing him with the excommunication with which Joshua banned Jericho and with the curse which Elisha cursed the boys and with all the castigations which are written in the Book of the Law. Cursed be he by day and cursed be he by night; cursed be he when he lies down and cursed be he when he rises up. Cursed be he when he goes out and cursed be he when he comes in…

Finally, people were recommended to keep away from him; the safety distance was four cubits at the time. However, the distance was not merely physical; all communication had to be avoided.

> (T3) [N]o one should communicate with him, neither in writing, nor accord him any favor nor stay with him under the same roof nor within four cubits in his vicinity; nor shall he read any treatise composed or written by him.

Three years later, in August 1659, two testimonies were presented to the Inquisition of Madrid in which a monk Tomas Solano y Robles and Captain Miguel Perez de Maltranilla testified that they had met on their travels "doctor Prado" (Daniel de Prado, Spinoza's friend) and "a certain Espinoza", who was "a good philosopher" (*buen filosofo*).[4] According to brother Tomas, the two were observant Jews (they refused some foods, for example), but they had been excommunicated:

> (T4) …because they thought that the Law was not true, and that the soul dies along with the body and that God exists only philosophically.

For many, the three claims that Spinoza had, according to these testimonies, denied, were crucial items of faith (see a, b and c below). For Spinoza, by contrast, they were open to free interpretation. He seems to have been aware that his views were risky but maintained that "the theologians' prejudices" were among the chief obstacles for people turning their minds to philosophy.

The first claim that Spinoza denied, according to the testimony of brother Tomas and Captain de Maltranilla, was that

---

[4] Kasher and Bidermann (1990).

the truth of the Law. This refers to denying the authority of the Torah, which, according to Maimonides, can happen in three different ways: (i) to say that the Torah is not of divine origin, (ii) to say that Moses said it himself, (iii) denying its interpretation in oral law.[5] In a word, the denial of the truth of the Law would mean denying its divine origin or its traditional oral interpretations.

Secondly, Spinoza was testified to maintain that the soul dies along with the body. This was a risky claim, since even others who took the soul to live after death but only for a short time were condemned as heretics. The problem is that if the soul dies with the body, there is no reward or punishment after death, which seems to undermine God'' authority. Menasseh Ben Israel'' important book *Nishmat Hayyim* (1652) that was published seven years before Spinoza'' excommunication and that Spinoza almost certainly owned,[6] also stresses the importance of soul'' immortality as the basis for reward and punishment. It has been argued that Spinoza was influenced by Alexander of Aphrodisias (around 200 CE) through Gersonides.[7] Alexander, in his turn, builds on Aristotle on the point that there is no personal immortality of the soul after the body dies. The contents of the intellect are eternal or atemporal and unchangeable, which means that some impersonal survival of the intellect is possible to the extent that one has managed to understand the intelligible structure of reality. However, if the immortality is impersonal and purely intellectual (concerning the immutable general structure of reality), it does not allow personal rewards or punishments. This seems to undermine the religious basis of morality. Therefore, Spinoza'' views about the soul'' (im)mortality most probably were important for the excommunication.[8]

As to the claim that God exists only philosophically, it is clearly not meant as a straightforward denial of God's existence. God is the only infinite substance Spinoza claims there to be, and it is identical with the whole of nature. However,

---

[5] Kasher and Bidermann (1990, 104).

[6] Adler (2014).

[7] Adler (2014).

[8] For the claim about soul's mortality as a central reason for the excommunication, see also Nadler (2001, 157-184).

to say that God exists "merely philosophically" can be understood as a denial of Providence, God's continued influence in nature and reality, that is also related to reward and punishment. Again, the grounds for God being the ultimate and unquestionable source of moral authority are at stake.

There is some discussion on whether Spinoza's actions were more important for the excommunication than his views. One instance of such action could be if Spinoza joined Menasseh Ben Israel's attempt to arrange Jews a possibility to travel to England.[9] However, it is plausible that merely expressing the kinds of views explained above, was a sufficient reason for the excommunication. The Jews seem to have wanted to project an image of being law-abiding citizens at that time in Amsterdam in order to protect their status in the city that was better compared to most of Europe. Therefore, Spinoza's focus on the freedom of thought and philosophical interpretation of religious doctrines, especially if he was taken to challenge the divine authority of the Torah, its oral interpretation and soul's post-mortal judgment, would most probably have been considered dangerous in and for the congregation.

2. "A Missionary's" Attitude to (Historical) Authors

Let us now turn to the examples of claims I have picked in order to illustrate the different attitudes we take to historical authors. My first example comes from Spinoza and the eighth proposition of the first part of his *Ethics* (1677). According to Spinoza:

> T5. [E]very substance is necessarily infinite (*Ethics* I, proposition 8).

This, however, is far from obvious. For instance, for someone who works on the ancient Greek philosophical tradition, a substance is the being (literally "beingness") for all things that are and claiming that there is only one infinite thing seems to be contradicted by experience. Focusing merely on who is right or wrong and building on something like philosophical instinct, one could object: "This is absurd! Not every

[9] Teichner (1957).

substance is infinite. Perhaps no substance is. Substances are (in) natural things and they are finite." From the perspective of Aristotle's *Metaphysics* and the later tradition building on it, the most important sources for the later discussion on substance and the origin of the term, one might continue: "It is an important reason for claiming that forms are substances that one needs to define natural things and explain the processes and changes in individual natural things. Therefore, there must be finite substances in nature."

From a Cartesian point of view, timewise close to Spinoza, the objection would be: "Bodies and minds or souls are substances and they are finite! Only God is an infinite substance." (To this Spinoza would perhaps respond: "Exactly!")

If we are only concerned with finding statements and theories that fit with our existent views and interests, there is probably not much for us to learn from Spinoza's *Ethics*. Those who work on the history of philosophy usually take this for granted: we are not primarily concerned with who is right and who is wrong.

Let us then consider another example of our possible agreement or disagreement with a past author. In the *Politics* (1.13), Aristotle says that, in women, the deliberative part of reason is without authority (*akuros*). He also states that some people are natural slaves (not necessarily those who actually are slaves) and they lack such a part altogether. Such groups of people – women and natural slaves – are thus better off being at least partly governed by another person's reason that has an authoritative deliberative part.

> (T6) The freeman rules over the slave after another manner from that in which the male rules over the female, or the man over the child; although the parts of the soul are present in all of them, they are present in different ways. For the slave lacks the deliberative faculty (*to bouleutikon*) altogether; the woman has it, but it is without authority (*akuron*), and the child has it, but it is immature (*ateles*). (*Pol*. 1.13, 1160a10-15)[10]

Most, if not all, of us probably disagree with Aristotle on these points. In contrast to the example about Spinoza's *Ethics*, it also seems that if I expressed agreement with Aristotle

[10] Jowett's translation from Nielsen (2015, 572).

(taken at face value), you would think that there is something wrong with my beliefs, values, and ethical/political judgments.

Again, if being right and fitting our expectations was all there is to philosophizing, a reader like this would leave both Spinoza's *Ethics* and Aristotle's *Politics* aside as sources that present implausible and/or unacceptable views. Let me name such an attitude as "the missionary". In no reference or disrespect to actual missionaries, I will just use the name here to express the attitude to a text that:

> (T7) "A missionary"
>
> (M1) focuses on whether the source agrees or disagrees with them or the general received views of their time.
>
> (M1a) if the source agrees, proceeds to promote the views (perhaps with special enthusiasm)
>
> (M1b) if the source disagrees with them or the received views, missionary readers proceed to discard the source and perhaps even consider its author a non-rational or otherwise substandard individual.

The core question in such an approach is whether a source agrees with what the reader already knows, the received views or some generally assumed common sensibilities.

Imagine, then, that I come across Jonathan Israel's arguments about Spinoza as a key thinker for enlightenment. Imagine also that I become convinced by Israel's arguments that Spinoza was a crucial figure for (Radical) Enlightenment for the following reason. According to Israel, Spinoza:

> (T8) …with his one-substance doctrine—that body and soul, matter and mind are not distinct substances but rather one single substance viewed under different aspects—extends this "revolutionary" tendency appreciably further metaphysically, politically, and as regards man's highest good, than do Descartes, Hobbes, or Bayle.[11]

[11] Israel (2009, 2).

According to Israel, this also means that Spinoza creates a sharp opposition between theology and philosophy, and this is what makes him a central figure of what Israel calls "the Radical Enlightenment".[12] To the extent that I am committed to the ideals of Enlightenment and perhaps its radical version, this perhaps creates some pressure for me to re-evaluate my views about Spinoza's theory of substance. If I am committed to the ideals of democracy and radical political equality of all humans and I think philosophy and philosophers should further those ideals, I should accept Spinoza's one-substance theory, since it is, according to Israel, necessary for supporting those ideals.

This raises the question whether such an attitude also falls under what I have called "the missionary" above. Perhaps there is a version of the attitude – we could call it "a political missionary" – that supposes that one should prioritize the (potential) political consequences of the theoretical philosophical views that one accepts and to choose one's other views accordingly.

(T9) "A political missionary"

(M2) Accept the theoretical philosophical views that best agree with your political convictions or the more general political convictions you detect in your historical period and society.

Assuming now that I accept (M2), I should thus go back to my earlier assessment of Spinoza's view on substance and revise my previous rejection of his one-substance theory.

It needs to be noted that although I am using Israel's argument to illustrate a point here, I tend to think that a one-substance doctrine could also give rise to a more off-putting ideal in which the individuals' political rights are subordinated to some goal that is common to all (for example, in a Hegelian or Kojèvean style that Eerik Lagerspetz and Sari Roman-Lagerspetz discuss in the collection).[13] It is also far from clear why a finite-substance theory could *not support* a diverse plu-

[12] *Ibid*. The Radical Enlightenment is opposed, according to Israel, with a moderate one. See Alexander (2020) for a three-fold division (radical, sceptical, liberal).

[13] Similar critique is also found, e.g., in Nietzsche, Kierkegaard and Deleuze.

ralist society in which individuals' political rights are of primary importance. These questions, however, have already created distance from the "missionary attitude". They take us beyond simple reactions of approval and denial to views presented.

It is also worth asking to what extent Spinoza himself might have accepted (M2). As has been argued by other scholars,[14] at the time when Spinoza was excommunicated, Jews in Amsterdam had to project an image of themselves as law-abiding citizens, and Spinoza's views endangered this project. The question then becomes whether Spinoza should have revised his views to agree with the important goals of the Jewish community in Amsterdam. However, he was perhaps more committed to the more general political ideals of equality, so that the local political goals of the community's self-preservation became secondary to those ideals.

My intention is not to deny that our theoretical views (e.g., in the area of metaphysics) have intended and non-intended political consequences. Although I am not sure I would agree on any particular formulation of this connection, it is not my purpose to deny its existence. What I would like to insist is intellectual honesty and responsibility in examining the theoretical (e.g., metaphysical) views as such and not primarily on the basis of some supposed political implications they may have.

3. Aristotle on women's authority

Leaving Spinoza for a moment now, let us go back to the passage from Aristotle's *Politics* 1.13 quoted above (T6). Although, at first sight, Aristotle seems to claim in the passage that the rational part of women's souls lacks authority, there are reasons to ask (i) whether he in fact subscribes to such a view and especially, (ii) if he does, whether this has something to do with the nature of women. Some scholars have argued that Aristotle in fact is *not saying* that women's reason is by nature without authority and that women should thus be subordinated to male rule and left without political power in the city. From this critical perspective, the point is that, at Aristotle's time, the women of Athens were, as contingent

[14] Kasher and Biderman (1990, 127) talk about "a low-keyed demeanour."

empirical circumstances had it, without authority because they lacked the kind of public role in which they could have exercised their reason's capacity to rule.[15] One version of this argument is that the arrangement still is, from Aristotle's point of view, beneficial because it functions for the goal of preserving households and the city.[16] The subordination of women thus becomes a combination of the ("natural") subordination of households to the political community and women's authority being (empirically) limited to the former. According to such a view, women's social role in the historical and contingent circumstances does not allow authoritative use (and perhaps training of the authoritative use of) their reason.

Other scholars have argued against the reading just outlined. Karen Nielsen calls it "the convention view"[17] and objects to it as follows. From Nielsen's perspective, Aristotle is committed to the claim that women are somehow naturally unfit to rule in a political community because women are *unfit to rule themselves* – much like akratics, who make good decisions but fail to carry them out. One of Nielsen's central arguments is the difference Aristotle makes between women and boys in (T6). If women merely lacked training of their deliberative capacities and that was the reason for their lack of authority, Nielsen argues, women should be in the same category with boys and their deliberative powers should be called "immature" or "imperfect" (*ateles*). However, this is not what Aristotle says but distinguishes women's deliberative reason from that of boys who merely lack training. Yet even Nielsen agrees with "the convention view" that there is no relevant *intellectual* difference between men and women, especially that men are no smarter than women. She notes that, in the passage in Aristotle's *History of Animals* (8.1, 608a28) on female softness implying lack of authority, women and females are in fact said to be better at learning than males.

Nielsen's argument has a plausible point to make about women and boys in (T6). There is, however, the following

---

[15] Saxonhouse (1982).

[16] See especially Deslauriers (2003).

[17] Nielsen (2015).

difference between boys and women: boys' deliberative reason is still in the process of growth and thus educable, while for women this is not necessarily the case any longer. If they have, as girls, missed their chance to train their authority, it could be difficult for them to learn it as grown-ups. There are also difficulties related to the parallel between women and slaves to which I will shortly return.

More generally, we need to ask to what extent such a "naturalistic" view of the political subordination of women agrees with Aristotle's view (*Posterior Analytics* 1.10, 76a31-40) according to which each science should have its own principles, unless a science is subordinated to another one as optics is to geometry or harmonics to arithmetic (*ibid*. 1.7, 75b12-17; 1.9, 76a22-25). There is no reason to assume without further argument that the science of politics would be subordinated to natural philosophy in a similar way as, e.g., optics is to geometry. Therefore, we need to ask how exactly principles and results concerning women and men in natural philosophy (including the study of soul) could be used in the science of politics (including ethics) in Aristotle's conception of sciences. Claims about nature also raise difficult questions about the role of nature in Aristotle's politics, e.g., what it means that the *polis* or the city state is something natural. Is it something that is natural in the sense of in some sense following from the human nature as the best possible organization of power relations and relations between people in which human beings can live? Or, is the *polis* rather natural in the sense of being an entity that exists by its own nature and what this means, given that cities do not grow from the earth.[18]

Some scholars have argued that Aristotle does not succeed in keeping his psychology separate from his political theory. This is because he notes that the soul's rule should be kingly (*basilikê*, *Politics* 1.5, 1254b5-6)[19] – a claim that resonates with Plato's remark that reason should rule in the soul like a king, not as a tyrant (*Republic*, 580c). Although for Plato, at least in the *Republic*, the isomorphism between the soul and the city is of central importance, there is no need to assume that Aristo-

[18] For one recent discussion, see Rapp (2021) and Horn (2021). Rapp defends a version of the former, Horn a version of the latter claim.

[19] Spelman (1983) quoted by Nielsen (2015, 574).

tle follows his teacher on this. Aristotle does not formulate a version of the isomorphism claim and perhaps just uses a similar way of speaking without literal commitment to the idea that the structure of the soul should be reflected in the structure of the ideal city.

One question related to the parallel between women and slaves concerns the so-called "natural slaves". Natural slaves in Aristotle are not necessarily those who in some historical and political circumstances *in fact are* slaves. Therefore, we need to ask how this distinction can be reconciled with the claim assumed in the readings above that "natural women" (i.e., those whose deliberative part is without authority whatever the reason for this is) are actual empirical women. This raises the further question about whether actual empirical men are "natural men", by essence fit to rule others. If we recognize the difference between natural slaves and actual empirical slaves, we should also allow a difference between natural women (or men) and actual, empirical women (or men), and similarly with girls and boys.

Finally, despite the apparent claim in (T6) that women should be governed by men, Aristotle stresses that the male rule over the female should be *political* rather than kingly, as father's rule over children is (*Pol*. 1.13, 1259a40-1259b1). The point about political rule in Aristotle is that, as opposed to the kingly one, it is a rule of *equals*. Therefore, whatever Aristotle says about the authority of women's deliberative capacity, women are assumed to be equal to men. One problem is that Aristotle uses an unhelpful comparison (as his comparisons usually are) to illustrate the relation between men and women. He says that:

> (T11) …when one rules and the other is ruled we endeavor to create a difference of outward forms and names and titles of respect, which may be illustrated by the saying of Amasis about his foot pan. The relation of the male to the female is always of this kind. (*Politics* 1.13, 1259b6-10)

The question is, of what kind exactly? Amasis (Ahmose II, 570-526 BCE) was an Egyptian pharaoh, who came from a modest background and, according to Herodotus, people failed to respect him at first. Then Amasis made an idol of

god out of his foot pan that people then started to idolize and transferred the same attitude to him.

At first sight, the comparison suggests that when there is rule of equals and the person who rules is without a noble background or other inherited credentials of power, external signs of power are especially useful. Therefore, the comparison supports the claim of equality between men and women and the special need (for males or females?) to create "outward forms and names and titles of respect" in order to rule. Nielsen argues against this reading by claiming that although at first sight the passage seems to support the convention view, the comparison underlines special executive powers that Amasis has: "[h]e knows how to push his will through".[20] However, I do not think this is what the passage (T11) is about. Rather, Amasis seems to be clever or resourceful despite his humble origins. If Aristotle had wanted to underline executive power, one would expect a comparison to a display of force. Moreover, if women are supposed to be weaker by nature than men (but cleverer), they are the ones that should be compared to Amasis in the anecdote. This entails that women can rule if they can display suitable external tokens of power.

In balance, I tend to think, that the "convention view" has more support in the passages. However, the main purpose of discussing the issue has been to show the sheer complexity of making sense of what Aristotle actually is saying about women and the authority of their deliberative power. While many of the passages fit "the convention view" very well, it is not evident how it can explain the comparison with boys in (T6). It needs to be said, however, that (T6) does not lend unquestionable support to the "nature view" either, as we can call views such as Nielsen's (women are naturally unfit to rule), and it also raises difficult questions about whether the men and woman, boys (and possibly girls) are empirical humans of a particular gender or sex.

4. Scholarly Concerns

[20] Nielsen (2015, 585).

I am using these texts here as examples to illustrate another approach that I call that of "a scholar". From the scholarly perspective, the question with respect to a text or a claim is not merely or primarily whether the source is right or wrong and whether it agrees with one's own views or a society's received views or not. From the scholarly perspective, authors are granted the benefit of the doubt and our task is to try to find out what they actually *are claiming* and how their arguments and claims should be understood. This enables but does not necessitate reflections concerning the plausibility of their views or the truth of the matter.

(T12) A scholar

(S1) Attempting primarily to make sense of what an author says.

(S2) Connecting what an author says to other elements in their philosophy.

(S3) Trying to make sense of what an author says in their philosophical- historical environment – possibly with respect to extra-historical or non-textual standards of plausibility.[21]

This list is not meant as a full methodological account of scholarly work in the history of philosophy. Rather, the point is to describe what kinds of questions a scholar necessarily focuses on in a given text and its historical-philosophical as well as non-textual context.[22] Straightforward "Enlighten-

[21] For non-textual standards of interpretation, see Kaukua and Lähteenmäki (2010). I use the notion of plausibility here that assumes some normative standards for evaluating what makes sense and what does not. For historical authors, the standards are largely related to the discussions of their time (what constitutes a sensible point in a given discussion). However, we also inevitably analyse plausibility from our own point of view. This is still different from assuming that we are in possession of a truth or of an inherently more plausible view than the authors of the past. Our perspective is also governed by contingent historical and philosophical circumstances, as Simo Knuuttila also emphasized; see his interview by Alhanen and Perhoniemi (2002).

[22] The approach I briefly describe here bears some resemblance to the one Michael Beaney calls "dialectical reconstruction' (2013) and describes very briefly in his helpful analysis of the history of rational reconstruction. I am grateful to Sara Heinämaa for drawing my attention to Beaney's article. See also Knuuttila on Hintikka (2006); Kaukua and Lähteenmäki (2010).

ment projects'[23] aside, scholars tend to assume today that the question of who is right and who is wrong – the question of truth about a given matter – is at least momentarily suspended. The primary task of a scholar is to make sense of a given text first from the perspective of what they see the author as trying to claim. Therefore, the main task of a reader of a historical text is to acquire understanding about a view that is, at least to an extent, alien. However, as Kaukua and Lähteenmäki underline, it is also important to identify common ground with historical authors in cases in which the shared assumptions are so obvious that they do not occur on the level of the text.

5. Concluding Reflections

One reason for me to have (briefly) discussed these issues here is the ideological pressure recently put on classics. For example, Princeton University announced in May 2021 that they will cancel their classics program in order "to combat institutional racism".[24] Elsewhere, the French minister of education wants to strengthen the teaching of Latin and Greek in the French *lycées* exactly to oppose "woke ideology".[25] One central question related to such attempts, not only within classics, is how we think that moral or political education works. Something analogous to the missionary approach is often found in Plato's *Republic*, in which moral education is envisioned in terms of stories with a clear moral, and even musical scales are banned if they arouse unwanted emotional reactions. Plato scholars today deny a literal or straightforward reading of these claims. If moral education worked in such a straightforward way – producing the qualities and

---

[23] In such Enlightenment projects, historical authors are treated as populating a lower stage of human development, poorly educated in the conceptual and theoretical refinements that the person approaching them is supposed to have access to thus granting them a position of superiority; for a description of such "Enlightenment approach", see Knuuttila's discussion on Hintikka's approach to history of philosophy (2006, 93-94).

[24] https://greekreporter.com/2021/09/06/greek-latin-princeton-sparks-debate-classics/

[25] https://www.letudiant.fr/lycee/jean-michel-blanquer-veut-developper-l-enseignement-du-latin-et-du-grec.html

dispositions of character in the youth that it is intended to produce – a Soviet (or DDR) education program would probably have been a great success.

It goes without saying that I am skeptical towards the idea that reading classics leads to definite political conclusions or supports any political agenda in any predictable way. Rather, in addition to the complexity of understanding the sources, I would like to stress the complexity of how people react to what they read and what kinds of effects certain moral/political education programs or arguments produce. There is no straightforward way of evaluating and/or manipulating the social and political outcomes of philosophical arguments and views. There is, of course, good reason to argue for the views that one takes as important and morally valuable, and philosophers can certainly contribute to social and political progress. (Consider, for instance, Mary Wollstonecraft's arguments for women's equal political rights.) However, in times when opinions abound as do black-and-white portrayals of any complex situation and even excommunications from the scholarly community are taking place, it is important to learn and teach to think for themselves – and to be able to see what one is doing oneself. Intellectual honesty is also called for. Ends do not justify the means.

Intuitions also differ with respect to how much we know at the moment. Past authors can remind us that we are tied to our own historical situation, and it is at least possible, if not probable, that we know rather little and that future generations will question many of the tenets we find obvious or unquestionable. There is also such a thing as healthy skepticism; carefulness and responsibility with respect to drawing firm conclusions and assuming that we have things figured out. By "healthy skepticism" I do not mean that we could never arrive at firm conclusions about what another person is saying or whether it is acceptable or not. Sometimes it is, of course, necessary to conclude that our conversation partner or a past author has views that are unacceptable from the point of view of today's moral and/or political sensibilities or our own views. However, this is compatible with keeping in mind that we as historical generations and individuals have our blind spots and that future generations or other individu-

als possibly find many of our historically, culturally and politically conditioned ways of thinking unacceptable.

In sum, past authors remind us how difficult it is to understand what an author or another person in fact is trying to say and what ends their arguments and claims are trying to promote – let alone what ends they actually promote(d). Past authors also help us to avoid not being smug about the progress humanity has made or the intellectual superiority we might tend to assume with respect to a historical author or another person. Moreover, there is no reason to buy Israel's suggestion that skepticism or epistemic humility should automatically lead to the promotion of merely moderate political change – let alone that skepticism necessarily comes with the extreme political conformism of someone like Sextus Empiricus. It is perfectly possible to combine healthy skepticism with important political goals. There are times in which resistance is needed and healthy skepticism is perfectly compatible with the ideals of resistance.

* * *

After my talk in the discussion, Simo Knuuttila asked about my own motivation to read the classics. I answered that my reasons are partly personal and partly more general. As to the personal, as a young student of philosophy, I found the epistemological debate about counterexamples to the definition of knowledge as justified true belief and the discussion around it as philosophically fruitless. Reading Aristotle, I was attracted to his way of discussing knowledge without being moved by (epistemological) skeptical challenges more extreme than those I would classify as being healthy. As to the more general reasons, it has been suggested that the real reason why so many central works from Greco-Roman antiquity went missing, were neither the fires in nor the sackings of the library of Alexandria. It was oblivion. Even if historical works exist in archives, libraries, or clouds but are not studied and discussed, they do not remain parts of our intellectual heritage if we do not actively engage with them.

Stockholm University

References

Adamson, Peter and Christoph Rapp, eds. (2021). *State and Nature: Studies in Ancient and Medieval Philosophy*. Berlin: de Gruyter.

Adler, Jacob (2014). "Mortality of the Soul from Alexander of Aphrodisias to Spinoza." In *Spinoza and Medieval Jewish Philosophy*, edited by Steven Nadler, 13–35.

Alexander, James (2020). "Radical, Sceptical and Liberal Enlightenment." *Journal of the Philosophy of History* 14, 257–283.

Alhanen, Kai and Tuukka Perhoniemi (2002). "Jumalasta, logiikasta ja rakkaudesta: Keskustelu Simo Knuuttilan kanssa." [On God, logic and love: A conversation with Simo Knuuttila.], *niin & näin* 2002/2, 57–61.

Auxier, Randall and Lewis Hahn eds. (2006). *The Philosophy of Jaakko Hintikka*. The Library of Living Philosophers 30. La Salle: Open Court.

Beaney, M. (2013). "Analytic Philosophy and History of Philosophy: The Development of the Idea of Rational Reconstruction." In *The Historical Turn in Analytic Philosophy*, edited by E. Reck, 231–260.

Deslauriers, Marguerite (2003). "Aristotle on the Virtues of Slaves and Women". *Oxford Studies in Ancient Philosophy* 25, 213–230.

Harding, Sandra and Merrill B. Hintikka eds. (1983). *Discovering Reality. Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*. Dordrecht: Kluwer.

Horn, Christoph (2021). "Normative Naturalism in Aristotle's Political Philosophy." In *State and Nature: Studies in Ancient and Medieval Philosoph*, edited by Adamson and Rapp, 59–81.

Israel, Jonathan (2009). *A Revolution of the Mind: Radical Enlightenment and the Origins of Modern Democracy*. Princeton: Princeton University Press.

Kasher, Asha and Shlomo Biderman (1990). "Why Was Baruch de Spinoza Excommunicated?" In *Sceptics, Millenarians and Jews* edited by Katz, 98–141.

Katz, David ed. (1990). *Sceptics, Millenarians and Jews*. Leiden: Brill.

Kaukua, Jari and Vili Lähteenmäki (2010). "Subjectivity as a Non-Textual Standard of Interpretation in the History of Philosophical Psychology." *History & Theory* 49/1, 21–37.

Knuuttila, Simo (2006). "Hintikka's View of the History of Philosophy." In *The Philosophy of Jaakko Hintikka*, edited by Auxier and Hahn, 87–105.

Nadler, Steven (2001). *Spinoza's Heresy: Immortality and the Jewish Mind*. Oxford: Oxford University Press.

Nadler, Steven ed. (2014). *Spinoza and Medieval Jewish Philosophy*. Cambridge: Cambridge University Press.

Nielsen, Karen Margrethe (2015). "The Constitution of the Soul: Aristotle on Lack of Deliberative Authority." *The Classical Quarterly* 65/2, 572–586.

Rapp, Christoph (2021). "Whose State? Whose Nature? How Aristotle's Polis is Natural." In *State and Nature: Studies in Ancient and Medieval Philosophy*, edited by Adamson and Rapp, 81–118.

Saxonhouse, Arlene (1982). "Family, Polity & Unity: Aristotle on Socrates' Community of Wives." *Polity* 15, 202–219.

Spelman, Elizabeth (1983). "Aristotle and the politicization of the soul." In *Discovering Reality. Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science*, edited by Harding and M.B. Hintikka, 17–29.

Teichner, Jacob L. (1957). "Why Was Spinoza Banned?" *The Menorah Journal* xlv, 41–60.

Articles

Alexandre Kojève and the Many Ends of History

EERIK LAGERSPETZ & SARI ROMAN-LAGERSPETZ

In 1992, American political scientist Francis Fukuyama published a book with an intriguing title: *The End of History and the Last Man*. The book was an extension of Fukuyama's earlier essay (1989). The fundamental idea, in the book as well as in the essay, was that the end of the Cold War was the end of history. Of course, there shall be all kinds of events even in the future. Technology will develop, the fashions will change, people will be interested in new things, and so on. However, the victory of liberal democracy, of capitalist market economy, and of the ideology of individualism they are grounded on, will be final and irreversible in the sense that they will not face any plausible ideological challenges.

Fukuyama's central thesis was generally rejected. The book and its title became symbols of the liberal euphoria that followed the Cold War and the collapse of the Soviet empire. The idea of "the end of history" was generally seen as a philosophical chimera that could be taken seriously only in the exceptional conditions of the early 1990's. Most people who use Fukuyama's title as a symbol of groundless optimism have probably not read the book, for the book itself is far more complex and ambiguous than its reputation. Fukuyama had a neo-conservative rather than a liberal background, and he himself was not fully happy with his conclusion. The second (usually omitted) part of the title of the book refers to Nietzsche and to "the last man", the truly despicable creature who appears in Nietzsche's *Also sprach Zarathustra*. Shadia B. Drury's early review of Fukuyama's work (Drury 1992–1993) provides a useful antidote to the usual reading; according to Drury, Fukuyama's fundamental thesis was "profoundly anti-liberal and anti-democratic" (80).

In his work, Fukuyama related his thesis of the end of history to G.W. F. Hegel, and, above all, to the Russian-French philosopher Alexandre Kojève. Kojève was one of the most influential interpreters of Hegel in the 20th century. In his famous lectures, held in 1933–1939 and published in 1947 with the title *Introduction à la Lecture de Hegel* (ed. Raymond Queneau, Kojève 1947/1969), Kojève argued that the end of history is a logical consequence of Hegel's philosophy – especially of the philosophy formulated in Hegel's *Phänomenologie des Geistes.* Kojève's interpretation was not a new one. Earlier, similar claims were made (very briefly) by Friedrich Nietzsche (1873–6/1997, 104–105) and by Friedrich Engels (1886/1973, 306–307). Both argued that although Hegel himself did not explicitly present the "end of history" thesis, it was an unavoidable consequence of his philosophy. For both, this was an argument *against* Hegel's philosophy, a sort of *reductio ad absurdum* result (Dale 2018, 22–78; Maurer 1965, 60). To Kojève, by contrast, the consequence was fully acceptable. Indeed, he considered it as the cornerstone of Hegel's system. History has an end, and humankind has already reached or is approaching that end. Kojève's attitude towards this supposed end of history is as ambivalent as Fukuyama's, but he never doubts that it is the correct way to read Hegel and expresses an important truth.

Kojève's – and Fukuyama's – interpretations are often rejected, not only as absurd but also as mistaken ways to read Hegel's philosophy.[1] Kojève himself admitted quite openly that his main purpose was not to present as loyal interpretation of Hegel as possible, but rather to use his interpretation as a way to express his own ideas in a Hegelian way. However, most Kojève's students and those writing under his influence have uncritically accepted his Hegel as the real Hegel. Our purpose here is not to take a definite stand for or against the fundamental thesis (Hegelian or not) of the end of history. Rather, we try to do the following. First, we consider briefly those less well-known thinkers who defended a similar thesis *before* Kojève. The thesis of the end of history may or may not

[1] Hegel himself talked explicitly about "the end of history" only once, in his *Introduction to the Lectures on the Philosophy of World History* (Maurer 1967, 446).

be genuinely Hegelian, but Kojève was not the first to argue that Hegel's theory implied the thesis, and that the thesis actually was true. Most thinkers who took the thesis seriously were, like Kojève himself, of Russian origin.

The thesis about the end of history has appeared in different arguments and in different contexts. It can be taken as a typical example of a philosophical idea without *prima facie* intuitive plausibility. For many people even the very talk about "the end of history" seems absurd. However, when looking at the thesis more closely, we are forced to admit that the unconvincing thesis follows from assumptions which themselves seem to be plausible. "The end of history" *seems* to provide an answer to problems that are genuine. Hence, we cannot simply sweep it aside without a further argument. Next, we say something about the reception and the influence of Kojève's interpretation of Hegel – for it has been extremely influential, especially in the post-war France. Finally, we will take up one important critique of the thesis.

End of history as a Russian philosophical theme

In this context, some clarifying comments are needed. The German "Ende", like the French "fin" and the English "end" are all ambiguous. They may refer either to a halting of a process, or to its purpose or aim. For example, the destruction of the human world would mean the end of history in the first sense, for after the event there would be no humans left. However, it would not mean that human history had achieved its aim or final purpose. In the latter sense, a process, for example human history, might have an "end" (in the sense of an aim or purpose) even if it never attains that end. *Kojève's "end of history" is an "end" in both senses of the term*. For Kojève, "history" is not just a temporal succession of events. A peculiar property of human history is that historical actors themselves are aware of the fact that they are part of that history, although they do not understand all the preceding conditions or the full consequences of their actions. They are motivated by (real or imagined) historical facts and use them as legitimation for action. The motivating and legitimating aspects of human history are related on one hand to the idea of *progress* (improvement, not just change), and on the other

hand, to the idea of *tradition* that makes actions parts of larger temporally defined wholes. Both ideas provide ways to justify actions. In a sense, Kojève's "end of history" means that both justifications become obsolete. According to Kojève, at the end of history, people are fully satisfied. They have neither a need to strive for something better, nor a need to legitimate the present by referring to the past.

Another Russian-French philosopher, Nikolai Berdyaev[2] has argued that "the end of history" is a specifically Russian theme (Berdyaev 1941/1957, 35). Besides Hegel, potentially relevant authors who have written extensively about this theme are a theologian and philosopher of religion, Vladimir Solovyov[3] (1853–1900), aforementioned Nikolai Berdyaev (1874–1948), and Moses Rubinstein (1878–1953). Rubinstein was born in Irkutsk but did his doctoral dissertation in Germany in 1906. These philosophers had several things in common. All of them were Russians and influenced by Hegel's philosophy of history.[4] They all discussed "the end of history" thesis before its most famous Russian-born proponent, Alexandre Kojève.

In his work *Three Conversations concerning War, Progress, and the End of History* (1900/1912, cf. Solovyof 1899/1918), Solovyov presents his version of our theme. The book consists of philosophical discussions, and the narrative about the end of history is embedded in them as a separate story.[5] In the story, Solovyov provides two versions of the end of history. One possible end of history is the unification of humankind by a secular idea of peace. For Solovyov, this secular (or degenerate Christian) version means the victory of the Antichrist: in it, human beings try to become divine by taking the place of God. The other, *real* end of history is that peace is reached when humankind universally accepts the Kingdom

---

[2] Transcribed variously as "Berdiaev", "Berdyayev" or "Berdjajev".

[3] Transcribed variously as "Solovjeff" or "Solovjev".

[4] Because President Putin has recently referred to Solovyov's and Berdyaev's philosophies of history in his speeches, it is important to stress that actually both thinkers rejected narrow-minded nationalism as a form of collective egoism and idolatry (see e.g., Solovyof 1899/1918, 295–298; Berdeyaev 1935/1961, 84–109).

[5] Dostoyevsky's *The Legend of the Grand Inquisitor* in his *The Brothers Karamazov* is an obvious model for this kind of narrative technique.

of God. Its earthly expression is a Christian world community. Human beings become divine in God by humbly accepting His will.

After immigrating to Germany, Kojève did his dissertation in Heidelberg on Solovyov's philosophy of religion (*Die religiöse Philosophie Wladimir Solowjew,* 1926). In the published version of the dissertation (Kojève 1934–1935), Kojève does not thematise Solovyov's version of the end of history thesis, but it is clear that he knew it. Kojève's relation to Solovyov has been studied systematically (e.g., Lowe 2018, chapter 3.), while his relation to Nikolai Berdyaev has not been studied at all. Nevertheless, the topic is potentially interesting. Although Berdyaev belonged to an earlier generation of Russian philosophers, he and Kojève had a lot in common. Both emigrated after the revolution and settled in Paris. Both moved between Orthodox Christianity and Marxism; while Kojève moved from Christianity to atheism, Berdyaev moved to the opposite direction. For both philosophers, Hegel-influenced philosophy of history and philosophical anthropology were recurrent themes. Both accepted a version of Existentialism before it came to fashion. They lived in Paris at the same time, and both were somewhat odd birds within the Russian émigré community. As far as we know, they never referred to each other in their works. We do not know whether there was any communication between them.

For Berdyaev, the end of history is a central theme (Berdyaev 1923/1990; Berdyaev 1941/1957). As for Solovyov, for Berdyaev it has a religious meaning. However, Berdyaev claims that Solovyov still sees the end of history as a *historical* event, as the beginning of an earthly Christian World-State. For Berdyaev, the end of history must itself be outside history (Berdyaev 1941/1957, 231–232). Although the necessity of the end is a part of the divine revelation, its (conditional) necessity is also revealed in the structure of our being. As he says, the end of history has at the same time an existential and a religious meaning (Berdyaev 1941/1957, 148).

Berdyaev's basic argument is the following: Either actions and events taking place in history have a (knowable) meaning, or they are meaningless. If the latter is true, our own lives are meaningless, too. If actions and events have a meaning, it must be related to their historical consequences, and, more

generally, to their place in the flow of history. Therefore, their true meaning can be known only when all of their consequences and their final place in history are known. This can be known only if history has a (knowable) end. If history has no (knowable) end, new events can always change the apparent meaning of the earlier events. Hence, either history has a knowable end, or all actions and events are without meaning.

Some critics of the traditional philosophies of history have argued in a somewhat similar way. For example, Arthur C. Danto (1965, 7–16; cf. also Löwith 1949, 5) argues that any historical event can have a meaning only within a particular context. In another context, it may acquire a different meaning. To take a simple example, at first most observers considered the Russian Revolution only a new bloody incident in the turbulent history of the country. When the Bolshevik regime stabilized itself, both the enemies and the supporters of the new government now saw the revolution as a world-historical turning point, either as the realization of a Utopia or as a mortal danger to the civilization. Now, after 30 years of the collapse of the regime, many historians have moved closer to the initial interpretation. By contrast, *human history itself*, as a complex event, cannot get its meaning from history. If human history as a whole has a meaning, it must come from outside of itself, from some ahistorical or trans-historical context (on this point, see also Sartre 1983/1992, 23). For Danto, this shows that the search for the meaning of history is futile, for all such contexts are dependent on controversial metaphysical or theological worldviews. Thus, Danto accepts the second horn of the dilemma in Berdyaev's basic argument: all historical actions and events are, in a sense, meaningless, because meanings ascribed to them are necessarily subjective or relative to particular contexts. While their meaning within a particular context may be fixed, it may change when we move from one context to another. Berdyaev cannot accept this; hence, he supposes that there must be a single transcendent context. It gives human history its meaning, and thus fixes the true meaning of every significant event in the process. However, he agrees with Danto in that the meaning of history cannot be found by studying the historical process itself. Ultimately, it is a matter of faith.

Berdyaev argues that in a contingent or eternally recurring historical process everything would be meaningless. If the world is totally contingent, every good or noble deed may become insignificant or evil in the future. In an eternally recurring world, all good deeds are futile, for they cannot make any lasting change. However, Berdyaev argues that even in a progressive view of history – central in the Enlightenment, in the early positivism, in Hegelianism, and in Marxism – every generation and its achievements can have only instrumental value, if progress is seen as an endless project. Endless progress is an inherently paradoxical idea, for it means endless sacrifices for the sake of a future that is never reached (Berdyaev 1941/1957, 229–230, Berdyaev 1923/1990, 180–182).[6] According to Berdyaev, this shows that if historical events are meaningful, history must have an end in *the both senses* of the term.

Kojève's end of history

Now we turn to Kojéve. We think that it is possible to distinguish at least three separate arguments for the end of history thesis. The first two are more or less implicit in Kojève's texts. Nevertheless, they are all interconnected.

The epistemological argument: The classical definition of knowledge is that knowledge is true justified belief. According to the traditional skeptical argument, for every contingent claim we take as knowledge, it is always possible that in the future new information or new arguments will show that the claim is not, after all, true or sufficiently justified. Hence, Kojève argues that either we do not have any knowledge about contingent matters at all, or there has to be a moment after which we will not acquire any new information or find new (valid) arguments (Kojève 1946/2007, 33). According to this epistemological argument, all philosophers who seriously claim that they have knowledge about something are committed to the further claim that, *at least in this issue*, history has reached its end.

---

[6] Beauvoir (1947/1976, 103–6) discusses the same problem. However, it seems that the paradox is avoided if steps towards the full realization of an ideal are *also* inherently valuable, not only valuable as a means to an end.

The ethical-political argument: We have supreme principles and ends (e.g., justice, freedom, universal happiness…). It is part of the notion of principles or ends that their full realization is possible. We can at least imagine a situation in which they are fully realized. When all principles and ends are fully realized, there is no need to strive for them anymore. However, human history *is* (largely) striving for unrealized principles and ends. Insofar as we have any supreme principles and ends, we have to admit that the end of history is at least conceivable.[7]

The epistemological and ethical-political arguments have a common ground. According to them, we human beings are teleological creatures. Our history consists of various attempts to attain knowledge or to realize principles and ends. The necessary starting point of this striving is that the aims we are striving for must be conceived as possible. Thus, having the idea of the final state seems to be constitutive for the human condition. At the same time, the idea contains a paradoxical element. If the desired final states were realized, there would be no further reasons for any significant action.

All the arguments sketched above (including Berdyaev's argument) are conditional: *if* history has a meaning, *if* genuine knowledge (in the classical sense) is possible, *if* we have supreme principles… then, the end of history is at least conceivable. In his *Introduction*, Kojève connects the arguments (sketched above) to his Hegelian view of the nature of society and of human history. There is no room to present Kojève's full theory here, but we may recall its general outlines. Kojève, as Hobbes, Hegel, and Marx before him, thinks that the society arises from conflicts, and that history is essentially conflictual. However, self-preservation and competition for scarce resources, emphasized by Hobbes, are not enough to explain why human beings are driven into conflicts. In the historical materialism usually ascribed to Marx, human history is a history of class struggles because human beings need products of labor to satisfy their material desires. In order to satisfy their needs, a process of production has to be organized, but effective organization of production unavoidably creates a division between the rulers and the ruled. Classes

[7] Cf. Sartre 1983/1992, 141: "All ethics presupposes the end of history".

struggle with each other for the distribution of the surplus created in the production process. Ultimately, they struggle for the control of the production process itself and thereby for the control of their own lives. Thus, the satisfaction of material needs is the motor of history.

Here, Kojève disagrees. Not all fundamental human needs are related to material satisfaction. Work is not the only distinguishing trait of human beings. Human beings desire "things" that are internally dependent on the desires, values, and actions of *other* human beings (Kojève 1947/1969, 6). They do not desire only material satisfaction, but also love and respect of others. In brief, they want to be *recognized*. If material scarcity were the only reason for conflicts, a world of abundance would necessarily be a peaceful one. By contrast, the desire for recognition cannot be satisfied simply by creating more of the desired "things". The desire for recognition is internally related to other human beings who have similar desires in the sense that they also desire love and respect. The desire for recognition, unlike desires directed towards material objects, is self-centered but, at the same time, *essentially* social. For Kojève – and according to his interpretation, for Hegel too – history is a struggle for recognition. Human beings, unlike other creatures, are willing to risk their lives for aims that are not related to their material survival (Kojève 1947/1969, 6, 4). Hence, Hobbes was wrong, and even Marx needs some correction.

For Kojève, the struggle for recognition leads to master-slave relationships. However, such relationships (and institutions built on such relationships) are necessarily unstable, for in such relationships neither the masters nor the slaves can be fully satisfied. A necessary requirement for recognition is that it is given *freely* (Kojève 1947/1969, 13, 21). Ultimately, the external signs of recognition given by slaves who are coerced or bribed do not satisfy masters' need for recognition. Fearful masses shouting slogans under the surveillance of the police, or courtiers with their flattering speeches, do not really satisfy tyrants' desire to be recognized. According to Hegel (and Kojève) a relationship in which people are treated as a means only (to use a Kantian expression) is psychologically unsatisfying and therefore unstable.

The third (explicitly Kojèvian) argument for the end of history is the following: The need for recognition is the fundamental motive for human action. Human history consists of human actions. The *universal homogeneous state* that recognizes all its members in an equal way and in which people recognize each other so that nobody is excluded from the social body of mutual recognition can satisfy the need for recognition of all human beings (Kojève 1947/1969, 44, 90, 95, 158–63, 236–7; Kojève 1946/2007, 42–44). Therefore, the arrival of the universal homogeneous state means the end of human history.

In Kojève's Hegelian view, this argument is connected to the earlier arguments. First, one central form of recognition is that the others are recognized as potentially relevant sources of knowledge. They are recognized as beings who have independent points of view that must be taken into account. Only when *all* points of view are taken into consideration, we can be sure that no important claims are left out. Thus universal (in Hegel's terminology, "absolute") knowledge and universal recognition are interconnected; universal knowledge is possible only in a state where everyone is equally recognized (Kojève 1947/1969, 32, 35, 95, 194). In this way, Kojève's recognition-based argument is related to the epistemological argument for the end of history. It is also related to the ethical-political argument. Because history is a struggle for recognition, all historically influential principles and aims must somehow be connected to that struggle. Only such principles are able to enjoy wide support. At the end of the day, only the principles of equality and general liberty can be successful, for they alone are compatible with universal recognition. Universal realization of the principles of the French revolution – liberty, equality, fraternity – realize the "universal homogeneous state" (Kojève 1947/1969, 50).

What is the nature of this "universal homogeneous state"? Kojève's description remains at an abstract level. He does not specify, for example, whether the universal homogeneous state would be capitalist or socialist. Kojève was a Marxist in his own idiosyncratic way, and in some of his texts, the universal homogeneous state could be identified with Marx's communist utopia. However, in other places he indicates that capitalist societies like the United States or Japan also realize

the end of history. In his posthumously published work on the philosophy of law (Kojève 1981/2000), the end of history is depicted as an individualist *Rechtsstaat*. As we saw, for Kojève's interpreter Fukuyama the end of history is unambiguously the final victory of liberal individualism. What is common for all these descriptions is that the end of history is a fully developed *modern* society. Kojéve is thoroughly and consciously a theorist of the modern world, and so, according to him, was Hegel. Nevertheless, Kojéve's formulation ("universal homogeneous state") sounds somewhat un-Hegelian. The modern state described in Hegel's *Rechtsphilosophie* is clearly a nation state characterized by internal complexity and differentiation (Roman-Lagerspetz 2009, 145). It is, perhaps, not an accident that Kojéve mostly ignores *Rechtsphilosophie* and concentrates on *Phänomenologie*.[8]

One important consequence of Kojève's interpretation is that Hegel's philosophy is itself, in two different ways, a part of the end of history it describes. According to the epistemological argument sketched above, all claims of knowledge imply that at least in some area of human knowledge, history has reached its end in the sense that no new (valid) arguments can refute those claims. However, Hegel is a special case, for he tries to understand the entire human history. According to his philosophy, history is essentially the development of collective self-understanding. Because Hegel is the first who recognized this, *his philosophy is itself a decisive step in the process it analyzes*. It refers to itself, and is its own proof, for nothing important can remain outside of it. It shows, not only its own possibility, but also its own necessity (Kojève 1947/1969, 49, 90, 94–96, 192–194; Kojève 1946/2007, 44–45). Its emergence itself shows that history is approaching to its final state.[9] At the same time, Hegel's philosophy is also *self-realizing*. When people become aware of the truth of Hegel's philosophy – for example, by reading Kojève's exposition of it – they start to act according to that philosophy, and by their own actions realize the universal homogeneous state (Kojève

[8] We are grateful to one of the referees for a useful comment.

[9] Kojève might happily accept Nietzsche's ironical comment that "for Hegel, the climax and terminus of the world-process coincided with his own existence in Berlin" (Nietzsche 1873–1876/1997, 104).

1947/1969, 98; Kojève 1946/2007, 56; cf. Roman-Lagerspetz 2022). Kojève's own later career as a bureaucrat who worked for the European economic integration could, perhaps, be seen as an example of this.

Critical arguments

It is difficult to overestimate the influence of Kojève's interpretation. It has been said that "the Hegel of the twentieth century is more or less Kojève's Hegel" (McGowan 2019, 139–140). Whole generation of French philosophers wrote under Kojève's spell. Jacques Lacan, Maurice Merleau-Ponty, Georges Bataille and Raymond Aron were among his students; Jean-Paul Sartre and Simone de Beauvoir, while not being among the regular students, were strongly influenced by his ideas (Butler 1987/1999; Roman-Lagerspetz 2009, 133). When the French philosophers started to criticize Hegel, the "Hegel" they attacked against was Kojève's Hegel. Philosophers as different as Beauvoir (1947/1976, 103–6, 116–8), Sartre (1983/1992, 88–90, 467–468), Albert Camus (1951/1987, 176–194), Emmanuel Levinas (1977/1998, 54, 79–80, 126) and André Glucksmann (1977/1980, chapter 3.) were largely in agreement in their critiques. They generally accepted the Hegelian-Kojèvian basic idea that the human history is a history of conflicts, and that these conflicts result from the universal desire for recognition. The human subject as well as the human society are arenas of these conflicts. However, these critics interpreted Hegel's philosophy as a philosophy of all-encompassing "Reason" that tries to overcome all conflicts and to suffocate all differences. Its logical conclusion would be the end of history as a totalitarian "tyranny of reason" (Drury 1994, 45). If the "end of history" is rejected, but the idea of history as a struggle for recognition accepted, what remains is the endless struggle without hope for redemption (Roman-Lagerspetz 2009). As Drury (1994, 76) says about Sartre: "Sartre thought that the best goals are those that are impossible. (…) The master-slave dialectic is a permanent feature of human existence." In these critiques, Kojève's Hegel, the supposed theorist of both the master-slave dialectic *and* the end of history, was self-evidently taken as the real Hegel.

As we noticed in the beginning of this article, the interpretation of Hegel as the philosopher of the end of history was not Kojève's invention. It was discussed already in the 19th century, and, as we saw, Nietzsche and Engels presented it (passingly) as an argument *against* Hegel. Interestingly, the Irkutskian philosopher Moses Rubinstein made an early critique of this interpretation in his dissertation (Rubinstein 1906). Rubinstein's arguments against the end of history thesis resemble the critiques put forth by the French philosophers some fifty years later; but unlike the French students of Kojève, Rubinstein tries to show that the thesis is not only false but also incompatible with Hegel's entire philosophy. Thus, Rubinstein's early dissertation could be interpreted as a pre-emptive reply to Kojève. Rubinstein remarks that for Hegel, the *aim* of history is unambiguously *human freedom* (Rubinstein 1906, 56, 65; Hegel 1837/1972, 54–55, 63, 104). In Rubinstein's interpretation, freedom in Hegel means the ability to act, and action means the possibility to create something new (new "objects of value", as Rubinstein puts it). In the supposed end (final state) of history, it would no more be possible to create something genuinely and significantly new. Hence, there could not be any freedom (Rubinstein 1906, 66). For this reason, Rubinstein argues, the interpretation of Hegel as a theorist of the end of history has to be mistaken (Rubinstein 1906, 62).

How would Kojève reply to this critique? Kojève (and following him, both Sartre and Beauvoir) states that the human subject is *negativity*. Human subjects are not constrained by any external limits. We are of course constrained by the physical world, but we have no such metaphysical, theological, biological or psychological essence that would force us to act in a predetermined way. Our freedom is our ability to overcome all such apparent predetermined limits, to negate any given definition of what it is to be a human being (cf. also Sartre 1983/1992, 69). Kojève presents this as a formula: "freedom = negativity = action = history" (Kojève 1947/1969, 209, 222, 226). For Kojève, freedom is unambiguously negative freedom. However, unlike the other theorists who talk about "negative freedom" he does not claim that freedom is just the *absence* of constraints (Lagerspetz 1998). Rather, freedom consists of action that *aims at removing or overcoming those*

constraints, an active negation (Kojève 1947/1969, 156). From this follows a paradoxical result: *When all the constraints of action are removed, there is no freedom*. To formulate the argument in a slightly different way: the moving force of all action is desire. According to Kojève, the specifically human desire is directed towards the other. Human beings, unlike other living creatures, desire recognition from the others, and are even willing to risk their own lives for it. This is the core of their freedom. Because, at the end of history, the human need for recognition is fully satisfied, there are no further reasons for this specifically human form of action (Kojève 1947/1969, 77). Hence, at the end of history, human beings are neither free nor unfree. Rather, they are like non-human animals whose behavior cannot be described in terms of free/unfree – when these terms are used in their Kojèvian sense (Kojève 1947/1969, 191–192, 220). When Rubinstein specifies that the thesis of the end of history is incompatible with Hegel's dialectical method (Rubinstein 1906, 61), Kojève's reply[10] would be that Hegel's *method* is not dialectical. Development through negations is not, as in the orthodox Marxism-Leninism, a universal philosophical principle, but a specifically human phenomenon. Dialectics is necessary for Hegel only as long as *the object of his study*, the human reality, is itself contradictory and develops through negations (Kojève 1947/1969, 179, 181, 191). Again, Kojève is fully endorsing a consequence initially presented by others as a *reductio ad absurdum*.

It is, however, clear that Hegel did not accept the Kojèvian-existentialist idea of freedom as pure negativity. In his *Philosophy of Right,* he connects the idea of freedom as the negation of all constraints to the terroristic period of the French revolution. Such a freedom cannot produce anything new (1820/1967, §5, 22). In Hegel, "negation" is always positive or productive in the sense that existing things are negated by creating new alternatives. The Hegelian "antithesis" is never just a denial or annihilation of something. Thus, Rubinstein

[10] Kojève never refers to Rubinstein's work, and we do not know whether he actually read it. In Kojève's *Introduction,* there are virtually no references at all to the earlier interpretations of Hegel.

rightly equates Hegelian freedom with *creation* (or "negation of negation") rather than with simple negation.

Conclusion

It seems that Rubinstein is right in his counter-critique. The idea of the end of history is incompatible with the fundamental thesis of Hegel's philosophy of history and political philosophy: the *aim* or *purpose* of history is freedom. It should be added that the textual support for this is much stronger than the support for the interpretation that Hegel would endorse the thesis of the end of history *as the final state*. This means that either Kojève (and his followers like Fukuyama) are wrong in their interpretation or, then, Hegel is internally inconsistent. The modern world cannot simultaneously be the realm of freedom and the end of history (when the "end" means the final state). We should notice that freedom as an "end" or aim is unlike other aims, say happiness or justice. A free society cannot be a static state of affairs, for freely acting human beings continuously change and transform it, and, following Rubinstein, produce new valuable things. However, a Hegelian might still claim that a free society is the aim of history without supposing that it can ever be fully realized; in Kantian terms, it would be a "regulative ideal" we could approximate endlessly.

However, we have not proved that Kojève's (or Fukuyama's) thesis of the end of history is meaningless as a possible interpretation of history. Moreover, we admit that even if Kojève's way to read Hegel is definitely one-sided (and sometimes even "violent", as some of his critics say) it also contains important insights. Sometimes one-sided interpretations may help us pay attention to aspects of classical works generally ignored in mainstream readings. Most notably, Kojève's fundamental insight that "recognition" is central to Hegel's whole philosophy, not just a single theme in his *Phenomenology,* is now generally accepted. Generally, we have to admit that Kojève's way to read Hegel, including his controversial and deeply problematic thesis about the end of history, has had a decisive impact on how many philosophers in the late 20th and early 21st century have understood not only Hegel but also the modern world.

University of Turku

References

de Beauvoir, Simone (1947/1976). *The Ethics of Ambiguity*. Translated by Bernard Fretchman. New York: Kensington Publishing.

Berdjajev, Nikolaj (1923/1990). *Historiens mening. Ett försök till en filosofi om det mänskliga ödet.* Translated by Stefan Borg. Skellefteå: Artos.

Berdyaev, Nicolas [Nikolaj Berdjajev] (1935/1961). *The Fate of the Man in the Modern World*. University of Michigan Press.

Berdyaev, Nicolas [Nikolaj Berdjajev] (1941/1957). *The Beginning and the End*. Translated by. R. M. French. New York: Harper & Row.

Butler, Judith (1987/1999). *Subjects of Desire. Hegelian Reflection in the Twentieth-Century France*. New York: Columbia University Press.

Camus, Albert (1951/1987). *Människans revolt*. Translated by Anders Ljungman. Lund: Alfabeta.

Dale, Eric Michael (2014). *Hegel, the End of History, and the Future*. Cambridge: Cambridge University Press.

Danto, Arthur C. (1965). *Analytical Philosophy of History*. Cambridge: Cambridge University Press.

Drury, Shadia B. (1992–1993). "The End of History and the World Order." *International Journal* 48, 80–99.

Drury, Shadia B. (1994). *Alexandre Kojève. The Roots of Postmodern Politics*. New York: St. Martin's Press.

Engels, Friedrich (1886/1973). "Ludwig Feuerbach ja klassisen saksalaisen filosofian loppu." In Karl Marx & Friedrich Engels, *Valitut teokset kolmessa osassa, vol 3*. 304–338.

Fukuyama, Francis (1989). "The End of History?" *The National Interest* No. 16, 3–18.

Fukuyama, Francis (1992). *The End of History and the Last Man*.

Fukuyama, Francis (1995). "Reflections on the End of History. Five Years Later." *History and Theory* 34(2) 27–43.

Glucksmann, André (1977/1980). *The Master Thinkers*. Translated by Brian Pearce, New York: Harper & Row.

Hegel, G.W.F. (1807/1972). *Phänomenologie des Geistes. Werke 3*. Frankfurt am Main: Suhrkamp.

Hegel, G.W. F. (1820/1967). *Philosophy of Right*. Translated by T. M. Knox. Oxford: Oxford University Press.

Hegel, G.W.F. (1837/1975). *Lectures on the Philosophy of World History. Introduction*. Translated by H.B. Nisbet. Cambridge: Cambridge University press.

Kojève Alexandre, (1947/1969) [1933–1939]. *Introduction to the Reading of Hegel. Lectures on the Phenomenology of Sprit, assembled by Raymond Queneau*. Translated by J. H. Nichols Jr. Ithaca: Cornell University Press.

Kojève Alexandre (1981/2000) [1943]. *Outline of a Phenomenology of Right*. Translated by B.-P. Frost & R. Howse. Lanham: Rowman and Littlefield.

Kojève, Alexandre (1946/2007). "Hegel, Marx ja kristinusko." In Alexandre Kojève, *Historian loppu*. Translated by Eetu Virén, edited by Tapani Kilpeläinen. Helsinki: Tutkijaliitto, 20–57.

Kojevnikoff [Kojève], Alexandre (1934) "La métaphysique religieuse de Vladimir Solovjev I." *Revue d'histoire et philosophique religieuses* 14, 534–554.

Kojevnikoff [Kojève], Alexandre (1935). "La métaphysique religieuse de Vladimir Solovjev II (suite et fin)." *Revue d'histoire et philosophique religieuses* 15, 110–152.

Lagerspetz, Eerik (1998). "Seitsemän vapauskäsitystä. Isaiah Berlinin 'positiivisen' ja 'negatiivisen' vapauden kritiikki." *Politiikka* 40, 87–104.

Levinas, Emmanuel (1977/1998). *Entre nous. On Thinking-of-the-Other*. Translated by Michael B. Smith & Barbara Harshaw. London: Atheleone Press.

Love, Jeff (2018). *The Black Circle. A Life of Alexandre Kojève.* New York: Columbia University Press.

Löwith, Karl (1949). *Meaning in History. The Theological Implications of the Philosophy of History*. Chicago: Chicago University Press.

McGowan, Todd (2019). *Emancipation after Hegel: Achievig a Contradictory Revolution*. New York: Columbia University Press.

Maurer, Reinhart Klemens (1965). *Hegel und das Ende der Gesichte. Interpretationen zur « Phänomenologie des Geistes »*. Stuttgart: Kohlhammer.

Maurer, Reinhart Klemens (1967). "Hegel et la fin d'histoire." *Archives de Philosophie* 30 (4), 483–518.

Nietzsche, Friedrich (1873–6/1997). *Untimely Meditations*. Translated by R.J. Hollingsdale. Cambridge: Cambridge University Press.

Roman-Lagerspetz, Sari (2009). *Striving for the Impossible. The Hegelian Background of Judith Butler.* Acta politica 36. Helsinki: Department of Political Science, University of Helsinki

Roman-Lagerspetz, Sari (2022). "Alexandre Kojève." In *Handbuch Anerkennung*, edited by L. Siep, H. Ikäheimo & M. Quante. Wiesbanden: Springer.

Rubinstein, Moses (1906). *Die logischen Grundlagen des Hegelschen Systems und das Ende der Geschichte.* Halle: Von Kammerer & Co.

Sartre, Jean-Paul (1983/1992) [1947–8]. *Notebooks for an Ethics*. Translated by David Pellauer. Chicago: Chicago University Press.

Solovyof [Solovyov], Vladimir (1899/1918). *The Justification of Good. An Essay on Moral Philosophy.* Translated by Nathalie A. Duddington. London: Costable & Co.

Solovjev [Solovyov], Vladimir (1886/1912). *Kolme keskustelua: sodasta, edistyksestä ja maailmanhistorian lopusta, sisältäen lyhyen kertomuksen antikristuksesta.* Translated by Toivo T. Kaila. Helsinki: Otava.

Pragmatic Truth and Realism in Pragmatist Philosophy of Historiography

SAMI PIHLSTRÖM

This essay builds upon my recent book developing a pragmatist philosophy of the humanities (Pihlström 2022). The philosophy of historiography is a central sub-field within the philosophy of the humanities, given the major role of history as a humanistic discipline. Neither in the book nor in this paper am I primarily investigating specific methodological questions of historiography (e.g., historical explanation); instead, I focus on the general realism issue as applied to the study of the past from a pragmatist perspective. In the 2022 book I try to argue that a pragmatist articulation of the realism issue, based on both classical pragmatists' such as Charles S. Peirce's, William James's, and John Dewey's ideas and their neopragmatist developments (by, e.g., Hilary Putnam, Joseph Margolis, and Morton White), provides us with a sophisticated conception of the value- and interest-laden ontology of humanistic inquiry, including historiography. Fortunately, pragmatist approaches to the philosophy of historiography have been defended in recent literature (e.g., Kuukkanen 2015; Gronda and Viola 2016; Kuukkanen et al. 2019; Viola 2020), and this is the discussion I hope to contribute to.

The following questions are, arguably, central in the philosophy of historiography as a sub-field of (pragmatist) philosophy of the humanities: Can historical truth be construed in terms of pragmatist conceptions of truth? In what sense can the pragmatist maintain that truth is what is satisfactory for us to believe (with reference to our future aims, values, interests, goals, and concerns) when the object of study is the past? Can the pragmatist claim the historical past to "really exist", given that the pragmatist conception of truth assesses truth-claims in terms of their (or their objects') future experiential significance? If we do endorse a pragmatist understanding of historiography, does this mean that writing history becomes indistinguishable from writing literature, as

Hayden White (1973) controversially suggested?[1] Does pragmatism commit us to an account of historical truth comparable to the view held by O'Brien, the Party torturer in George Orwell's *Nineteen Eighty-Four* (1949), who maintains that truth depends on the opinion of the Party?

After having introduced the issue of historical truth from a pragmatist perspective, taking my lead from questions such as these, I will discuss specific aspects of pragmatist philosophy of historiography, including the Peircean account of "real generals" and holistic pragmatism.

Historical truth

One of the classes of truths considered in Orwell's famous novel is, indeed, the class of historical truths. Winston, the protagonist of *Nineteen Eighty-Four*, claims to remember that there were airplanes before the Party came to power. O'Brien, however, denies this: if the Party claims that there were no airplanes before the rule of the Party, then there were no airplanes before the rule of the Party. Truth, according to O'Brien (and the Party), is just what the Party claims to be true; moreover, this "theory" of truth itself is, at the meta-level, something that the Party claims to be true, too. This thoroughgoing historical contingency – what the Party contingently states is the case – creates a *metaphysical instability* to the Orwellian world, as the opinion of the Party could randomly change. Any truths or facts would change along with such changes. Presumably even the view the Party holds about what truth *is* could change randomly. Even if the Party suddenly declared that it no longer maintains that truth depends on its opinion, *this* allegedly more realistic and objective account of truth would depend on its authority. When a dictator starts to meddle with truth, there is no easy return to any ordinary understanding of truth; indeed, truth, or at least

[1] White's (1973), Ankersmit's (2009), and other "narrativist" philosophers' of historiography, controversial views have raised a considerable debate whose core, arguably, is the issue of realism. One of the merits of pragmatism in this area is the possibility of critically integrating (non-metaphysical) realism and (some form of) ontologically constructivism. As usual, pragmatism seeks a critical middle ground here – as we will shortly see.

the availability of the concept of truth in our lives, can be genuinely destroyed, as in *Nineteen Eighty-Four*.

The totalitarianism imagined by Orwell is more radical than any real-life totalitarianism has ever been. For example, in the Soviet Union, it was (as it is in today's Russia and China) crucial for the rulers to control what people believe, hiding the "real" truth by, e.g., manipulating photographs (which, of course, is easier today than in the days of the Soviet dictators). This, however, presupposes that (historical) truth remains important: truth is respected in the attempt to hide it. In the Orwellian context, however, the Party's use of extreme violence succeeds in abolishing any distinction between what is true and what the Party claims to be true – to the extent that Winston, when horribly tortured, comes to believe that 2+2=5.[2] Similarly, historical truths about, say, airplanes, or anything else, are constituted by what the Party says, and there is no external control for checking its opinion against any objective fact.[3]

Orwell's novel has been interpreted as a defense of realism and the objectivity of truth – also by pragmatists (e.g., Mounce 1997). In contrast, Richard Rorty's (1989) controversial reading of *Nineteen Eighty-Four* emphasizes the concept of *freedom* in contrast to the concept of truth.[4] Mediating between these positions, a pragmatist account of the humanities may successfully counter the absurd reduction of pragmatist conceptions of truth to O'Brien's antirealism while acknowledging the dependence of historical truth on our (future-oriented) value-laden perspectives. A plausible starting point for such a pragmatist view can be found in William James – particularly in *The Meaning of Truth* (1978 [1909]), where he

---

[2] How firmly the novel is actually committed to this may be a matter of some interpretive debate, though.

[3] This is not to say that dictators like the leaders of Russia and China would not be very dangerous to human civilization in their disrespect for both truth and fundamental human rights, though Orwell's Big Brother is even more extreme. Yet, liberal democratic society does not *guarantee* that truth will prevail, either; as we know, truth has been in real trouble with democratically elected post-truth leaders, such as Donald Trump in the United States.

[4] See Conant 2000 for an insightful critique. Cf. also Kivistö and Pihlström 2016, chapter 5, for explorations of Rorty, Orwell, truth, and truthfulness.

seeks to show that pragmatism can very well accommodate a robust conception of historical truth, and of the existence of historical persons and events.[5]

Truth is *pragmatically needed* in historiography, especially in order for us to make sense of historiographical *disagreement* as an element of our historiographical practices. If there were no truth to be pursued and possibly discovered (however fallibly and insecurely), there would be no point in disagreeing with anyone else's interpretation of history; inquiries would collapse to mere subjective preferences. This, however, is not to believe naively that historians would primarily disagree about whether some particular event took place. Usually (as constructivist views on historiography insist) their disagreements concern broader interpretations of history, instead of straightforward "facts". Yet, the very idea of disagreement – even about plausible interpretations of whatever facts there are – presupposes that we are engaged in the practice of pursuing truth and objectivity. A pragmatist conception of truth is sufficient for the relevant kind of truth-aptness invoked here: historiographical constructions do not refer to ready-made facts "out there" in the metaphysically given historical past but to complex historical realities partly constituted by historians' value-laden practices of interpretation.

Truth vs. warranted assertion

Pragmatists have fortunately not simply agreed with Rorty's deflationary views but have developed a variety of accounts of truth and realism in the philosophy of historiography, ranging from historicist antirealist "narrativism" to Peircean realism and "rational pragmatism".[6] Jouni-Matti

[5] In cases of "historic truth", "it is but one portion of our belief reacting on another so as to yield the most satisfactory total state of mind" (James 1978 [1909], 54). For James's discussion of the real existence of Julius Caesar as compatible with the idea of true statements having "functional workings", see ibid., 120-122 (cf. also 150-151).

[6] While one might suppose that pragmatists generally defend antirealist views such as Hayden White's, relatively few actually seem to do so. Peircean realism about historical knowledge has been developed by many recent contributors to pragmatist philosophy of historiography (e.g., Colapietro 2016; Topa 2016; Laas 2016), while others have been inspired

Kuukkanen's (2015, 2017) "postnarrativism" deserves brief comments here as a major recent contribution to pragmatist philosophy of historiography. My views are in many ways close to his, especially regarding the need to synthesize realism and constructivism (or objectivity and subjectivity) and the understanding of historiography as a rational practice of inquiry. Kuukkanen plausibly develops these ideas by critically moving beyond the (partially correct) narrativist insights of Hayden White and Frank Ankersmit, and by arguing that historiography must be considered both subjective and objective at the same time (Kuukkanen 2015, 175, 200). While antirealists are right to maintain that historiography does not correspond to any definite "entities in the past" in the sense of "mirroring" historical reality (ibid., 9),[7] it does not follow that historiography cannot be distinguished from fiction-writing.

However, Kuukkanen's pragmatism, though rightly opposed to antirealism, remains unstable in its attempt to avoid claiming that historians are in the business of discovering truths (e.g., Kuukkanen 2015, 11, 2017, 90) and in suggesting that truth can be replaced by "warranted assertion" (Kuukkanen 2015, chapter 7, 2017, 95). He proposes to "reject absolute truth-functional standards and replace them with a cognitively authoritative rational evaluation without implying that there are absolutely correct interpretations" (Kuukkanen 2015, 2). While historiography is a discursive and argumentative practice of reasoning for or against theses (ibid., 66-67, 198-199), and hence a rational practice answerable to critical normative standards (ibid., chapter 5), this does not, Kuukkanen argues, make it a search for the truth.

---

by, e.g., James's, Dewey's, C.I. Lewis's, and W.V. Quine's pragmatisms (see essays in Gronda and Viola 2016; Kuukkanen et al. 2019). For a detailed investigation of Peirce's views on history, see Viola 2020. Viola's book is presumably the most substantial study so far on a pragmatist classic's (Peirce's) conception of history, emphasizing the various sources and dimensions of Peirce's ideas (viz., history of science, history of philosophy, semiotics) and particularly focusing on his realism about history.

[7] Kuukkanen's (2015, chapter 3) elaborations on constructivism as one of the central tenets of narrativism, critically retained in postnarrativism, is obviously relevant to pragmatist forms of constructivism as well.

His main argument against the truth-functionality of historiography relies on the premise that no historical entities in the past exist as truth-makers for historical interpretations (ibid., 11). Pragmatist "redefinitions" of the meaning of truth are not helpful, either (ibid., 138-142); Kuukkanen finds it more promising to abandon the view that historiography aims at truth, replacing this idea with the Deweyan notion of "warranted assertion" (ibid., 143ff.). However, the problem with truth-makers can be resolved by embracing a richer pragmatist conception of the historical past as ontologically dependent on historians' value-guided selections of relevant facts.[8] Combining pragmatically understood truth with an ontologically serious *pragmatic realism* helps us accept that historians do engage in the practices of pursuing the truth. Nor do I find it necessary to reject what is sound in "representationalism", another philosophical interpretation of historiography criticized by Kuukkanen (ibid., chapter 4). A relatively ordinary pragmatic notion of a historical interpretation being "about" something "real" in the past is sufficient; no specifically historiographical "non-representationalism" is needed to reject the implausible idea that historiographical writing refers to "unique corresponding entities" (cf. ibid., 64).

In addition to Kuukkanen, a balanced pragmatist response to the challenges of realism, truth, and objectivity in the philosophy of historiography has been formulated by Marek Tamm (2014). He is concerned not as much with applying theories of truth to historiography as with the question of which conceptions of truth and objectivity are "practiced within the discipline of history" (ibid., 266). Utilizing the theory of speech-acts (like Kuukkanen 2015, 158-161), he analyzes this in terms of what he calls a "pragmatic 'truth-pact'": "the conditions of historical truth depend on the illocutionary

[8] Some details to this general argument will be provided below in the context of "real generals" and holistic pragmatism (see further Pihlström 2022, chapter 3). In addition, while agreeing with Kuukkanen (2017, 114) about the primarily cognitive function of historiography, I would insist on the irreducibility of its ethical function. Even when criticizing Kuukkanen's rejection of truth, I find his theory of justification in historiography plausible (Kuukkanen 2015, chapter 9) – my only reservation being that it needs a (pragmatist) concept of truth.

force of historical utterance", and this "pact" is "'guaranteed' by fellow historians" and thus by a "disciplinary consensus as to the methods of inquiry, cognitive values and epistemic virtues" (Tamm 2014, 271). Truth thus becomes an epistemic notion, as the "truth-pact" is based on a "critical analysis of the available evidence" (ibid.).

The "truth-pact" is an implicit agreement between the historian and their addressees, according to which the historian's intention is to tell the truth. It presupposes "rules" – normative principles structuring our practice of pursuing the truth – according to which, for instance, the historian making historiographical assertions must be able to provide evidence and reasons for their truth-claims, and the asserted propositions must not be "obviously true". The historian must also be sincere in the sense of believing in the truth of their assertions. (ibid., 274–275). Following such rules obviously does not guarantee truth, but these normative principles "form the necessary grounds for a historical assertion to be considered as true" (ibid., 275). They could, therefore, be viewed as conditions for the possibility of historical truth – conditions based on how historiographical practices function.

One of the implications of Tamm's view is that historiography must be carefully distinguished from fictional writing that does not purport to be true (ibid., 277). Hence, narrativist philosophy of historiography must be firmly rejected.[9] On the other hand, the "truth-pact" does not, according to Tamm, involve any correspondence theory, as any supposed correspondence between the historian's claim and the historical fact would be "impossible to check" (ibid., 278). The critic may ask why this would be any easier to "check" in the natu-

[9] In this sense, Tamm's position is shared by Kuukkanen (2015, 2017), but the difference between the two is that Tamm believes we need to employ the concept of truth – albeit not a non-epistemic or justification-transcendent one. Generalizing their arguments, we may suggest that antirealists are wrong in supposing that our having to invoke ethical and aesthetic criteria when "selecting" (value-laden) historical facts for interpretation precludes epistemological and ontological considerations of such value-laden factuality. For the pragmatist philosopher of the humanities, aesthetic and ethical values are, while always present in our humanistic inquiry, entangled with ontology and epistemology. (See again Pihlström 2022.)

ral sciences, or in other practices of truth-seeking. Why, then, wouldn't Tamm's rules, or the "truth-pact", be applicable within *any* practice of inquiry pursuing the truth? Accordingly, while I find Tamm's discussion lucid and helpful – also agreeing with his pragmatic analysis of objectivity as "a way of *doing* things" (ibid., 281), a practice-embedded epistemic virtue rather than a mysterious property of the objects of inquiry – I am not convinced that his epistemic construal of truth crucially helps us in capturing the core of historiography as a pursuit of truth. I would be happy to retain most of what he says about truth and objectivity as practical pursuits within our fallible practices of historiographical inquiry, while grounding the pursuit of truth in the *ontology* of historiography – understanding that ontology itself as irreducibly value-laden. Truth will then fall in its place as a pragmatist "species of good".

Although I cannot go into details here, it may be suggested that the account of truth that my more ontologically oriented pragmatist philosophy of historiography employs operates with a notion of truth that is more strongly "justification-transcendent" than either Kuukkanen's (who basically replaces truth by warranted assertibility) or Tamm's (who maintains truth but defines it epistemically). I would be prepared to argue (though not here) that the pragmatist needs, for pragmatic reasons, a realist conception of objective truth that may transcend our contingent capacities of justifying our beliefs. There may, for instance, be "buried secrets" in history, that is, historical truths that will never be known – and recognizing this is part of our pragmatically justified ontology of the historical past. In brief, one does not have to give up pragmatism in order to be a moderate realist, nor vice versa.

Real generals

Therefore, we need to take a closer look at pragmatist ontology of historiography. It would be impossible to any compre-

hensive account of such ontology here, but I will offer one suggestion.

Historical inquiries do not merely refer to concrete particulars, such as unique historical events. They may also take more abstract entities and phenomena among their objects, such as (for example) historical tendencies of development and interpretive possibilities. These can be analyzed in terms of Peirce's realism about "real generals" (e.g., habits, dispositions, laws), which has rarely, if ever, been developed in the philosophy of the humanities in any detail.[10] We may argue that such generalities can be, and often are, among the ontological postulations of historiographical interpretation. Without a proper account of generality, our conception of the ontology of the humanities remains inadequate.

This seemingly abstract philosophical issue is clearly relevant to methodological questions of historical explanation. Historians explain why a certain event took place at a certain time, and explaining this typically involves explaining why some *other* event(s), among the countless events that *could* have taken place, did *not* take place; therefore, counterfactual questions ("what if…?") and contrastive counterfactual explanations ("why did X happen, instead of Y?") are part and parcel of historical scholarship and have been thoroughly explored by philosophers of historiography. Thus, the historian does not merely seek to find out what exactly happened ("*wie es eigentlich gewesen*") and why. Their focus is on the fact that something happened *instead of* something else.[11] This fact

---

[10] For relevant writings by Peirce, see Peirce 1992–98, especially vol. 2 (the theme runs through Peirce's writings as a whole, from the 1871 "Berkeley review" to his late essays on pragmatism and pragmaticism). For relevant *loci*, see Peirce 1931-58, 1.15–26, 4.1ff., 5.59–65, 5.93–101, 5.312, 5.423, 5.430–433, 5.453ff., 5.470, 5.502–504, 5.528, 8.7–38. For my more Jamesian pragmatist appropriation of Peircean realism, see Pihlström 2009. Viola (2020) emphasizes Peirce's historical realism; my use of the concept of "real generals", though inspired by Peirce, aims at no scholarly accuracy as an interpretation of Peirce. We should feel free to employ this concept for our purposes here without necessarily subscribing to the overall position Peirce labeled "extreme scholastic realism".

[11] See several essays in Tucker 2009. That is, both contrastive historical explanations – explaining why something happened instead of something else that could have happened – and the more popular genre of writing

must be taken not only *explanatorily* seriously (in terms of contrastive and counterfactual explanations) but also *ontologically* seriously; it is a fact requiring interpretation and thus the entire value- and purpose-oriented scheme of inquiry through which we conceptualize historical reality. In this sense, the ontological status of historical facts as "generals" (in addition to their being in an obvious sense "particulars") is based on their value-ladenness. It is only in the context of our value-based selection that they (can) ontologically emerge as the general facts they are.[12]

A simple example can illustrate this. A war historian may ask not only why Finland ended up in two wars against the Soviet Union (i.e., the Winter War in 1939–1940 and the Continuation War in 1941–1944) during World War II – in contrast to the counterfactual possibilities of having ended up in no war at all or, say, only one. They may also ask, and have asked, whether Finland could have avoided these wars. Clearly, if questions about historical possibility such as this are among the kind of questions asked by historians, history is not merely concerned with what actually happened. It is irreducibly concerned with modally complex questions about what happened in the context of possibilities that failed to actualize, and these may include specific questions about how exactly, say, the Continuation War might have been avoided (or not). Questions like these suggest that historians investigate *a modally complex reality*. Even if they do not primarily view their efforts as aiming at the truth but only at a plausible interpretation with some degree of objectivity and support by evidence (see again Tamm 2014; Kuukkanen 2015), in the absence of the pragmatically realist idea of a modally complex world of "real generals" being "there" – albeit available only through our practices of inquiry – the historians' practice would be incomprehensible. The norma-

---

"counterfactual histories", i.e., stories about how things might have developed had they not developed the way they did, arguably need Peircean realism of generals.

[12] On the value-ladenness of history, see also, e.g., Dray 1980, 28, 42–46. As Dray puts it, when we recognize this value-ladenness, "the past as it actually was" will "coincide" with "the past as it must appear from the standpoint of a certain scheme of values – political, aesthetic, social, moral, intellectual, and so on" (ibid., 46).

tive context of inquiry within which it is *possible* to ask general questions about possibilities presupposes that historical inquiry is not just a matter of telling the most plausible-sounding story (compatible with relevant evidence) but a fallible practice of seeking and discovering truths about the real world and its real possibilities.

It is, therefore, we may argue, only within a Peircean realism of generality that we can make sense of the idea that while only actual historical events really "exist" in the past (to the extent that past events can be said to "exist" in some sense), historical scholarship also studies real generals, such as non-actualised historical possibilities, i.e., things that could have existed but did not and do not. Thus, it studies the reality of what did exist in the context of what could have existed.

Another context within which historians arguably need the ontology of real generals is their use of "colligatory concepts" (Kuukkanen 2015, chapter 6). By employing such synthesizing concepts, the historian interpretively organizes the past and makes it intelligible. The "antirepresentationalist" maintains that historiography is not referential or representational as it employs such concepts that do not refer to any corresponding entities in historical reality; for example, the concept of "renaissance" does not represent any clearly identifiable object or event in the past. The proposal to interpret historical ontology in terms of real generals resolves this problem by making our ontology rich enough to provide reference, and even truth-makers, for colligatory concept-use and theorisation.[13] Whatever colligation exactly amounts to, the pragmatist may suggest that its historiographical referentiality can be accounted for in terms of pragmatic realism ontologically accommodating "real generals". Even though there are no "ready-made" entities "out there" in the historical past just waiting to be picked out by colligatory concepts (as if renaissance would just exist independently of the use of

[13] In his admirable discussion of colligation, Kuukkanen (e.g., 2015, 106, 113) is somewhat unclear in criticizing the idea that colligatory concepts would be "true"; truth should not be attributed to concepts but only to statements or theories. Again, a pragmatist account of truth (and of real generals) enables us to make sense of colligatory concept-use while preserving the idea of historiography as truth-seeking. Colligation itself is in the business of truth, pragmatically understood.

the concept of renaissance), our use of such concepts is not ontologically neutral, because it takes place within a practice of inquiry seeking to represent a modally complex world.

Holistic pragmatism

A specific version of pragmatism that I find highly fruitful for developing a plausible philosophy of the humanities is the *holistic pragmatism* proposed by Morton White (e.g., 2002, 2005).[14] Holistic pragmatism is not restricted to the humanities but functions as a general theory of inquiry; yet, it is particularly well suited for a critical pragmatist philosophy of the humanities especially because it enables us to transcend the boundary between *the factual and the normative*. This, arguably, is essential in historical scholarship, too, both in formulating research hypotheses and in their critical evaluation. According to White, our beliefs – factual and normative beliefs included – constitute a holistic totality: whenever testing any of them, we are testing their conjunction as a whole.

White (2002, 79) seeks to preserve a form of historical realism in the sense that the historian aspires to "tell the truth as natural and social scientists do", although the primary aim is not generalization but (usually) the construction of a historical, and often explanatory, narrative, with the *explanandum* and the *explanans* both guided by human *interests* (see ibid., 89ff.).[15] Following the classical pragmatists, White rejects any "picture or copy view" of history, as firmly as any picture or

---

[14] White's earliest publications on holistic pragmatism are from the 1950s (as are some of his early writings on the philosophy of history; see White 2005), but I am here only citing his late formulations of this idea. In some recent work (especially Pihlström 2021), I further develop holistic pragmatism in the philosophy of religion, in particular. Kuukkanen (2015, chapter 2) only discusses White as one of the "early narrativists", without paying attention to his later pragmatism.

[15] White's discussion of the role of interests in historical explanation is highly relevant to a pragmatist theory of historiography, but the holistic-pragmatist idea of integrating factual and normative components within our belief system tested as a totality is even more important – though it also entails a version of "historical relativism" that needs to be critically assessed in terms of holistic pragmatism itself (for a critical discussion, cf. White 2005, chapter 6).

copy view of science (ibid., 101). No truths, as James (1975 [1907], especially Lecture VI) reminded us, can be usefully regarded as mere copies of an independent reality. White (2002, 106) refers to the *plurality* of interests at work in historiography, as constrained by the "ethical rules of the historical game":

> It seems to me that we should recognize that the interests of the historian determine (a) his *terminus a quo* and his *terminus ad quem*, (b) the entrenched and the major intervening items in his chronicle, and (c) the items he selects as causes from his varying points of view. In that case, we should also recognize that his construction of a history is determined not only by the historian's interest in telling nothing but the truth but also by his other interests. (Ibid., 102.)

Our interests are, inevitably, *value-laden*. This makes even the historian's causal claims value-laden: "although a historian tries to present causal truths about some subject in a narrative, he chooses *what* truths to link causally on the basis of value judgments that a fellow historian may not accept when telling a true story about the same subject" (White 2005, 40; see 49–50). White rejects radical relativism, however, as well as a straightforwardly pragmatist view he associates with James; we cannot, he maintains, "do" anything with the "dead and buried" historical past with regard to our "practical aims" (ibid., 45–46). Here I would propose a more subtle understanding of "practical aims": even the most antiquarian historical interest must, according to (Jamesian) pragmatism, serve possible future experience. Historians' interests need not be practical or forward-looking in any immediate sense, but there must be a dimension of future-orientation in them. Otherwise, we cannot even find any pragmatic meaning in the historical concepts we use. After all, according to Peirce's and James's Pragmatic Maxim, it is in terms of the conceivable practical effects of (the objects of) our concepts and conceptions – including the objects of our historical concept(ion)s – that we must determine what those concepts and conceptions mean for us.[16] The value-guided future-orientation thus

[16] See the canonical formulations in Peirce 1992–98, vol. 1, and James 1975 [1907], Lecture II. In this regard, Tamm's (2014) pragmatist philosophy of

pertains to the meaning of historiographical concept(ion)s and interpretations, not simply to their truth.

White formulates a clear and, in my view, generally plausible account of the entanglement of truth and interests as his conclusion:

> [A] historian's total judgment of what should be recorded in a history […] is a blend of two judgments: one of truth and one of memorability. Truth may be the stronger claimant and the more objective, but memorability is always a factor in spite of being subjective. The absence of an objective criterion of historical importance makes it very difficult to show that historical truths about a society's economy or politics are more memorable than those about its philosophy. It also shows that there are more historical truths than are dreamt of by historians. (Ibid., 50.)

Accordingly, the holistic pragmatist admits that value judgments about what is more or less "memorable" or historically interesting play a key role in the historian's choice of "facts" – even causal facts – to be included in a historical account of any phenomenon to be explained. Thus, values are inescapably relevant along with "mere facts" not only in the methodology of historiography but in its ontology, too.[17] Better, there are no value-independent "mere facts" at all. Historical facts are what makes our historiographical statements true (if they are true), but those very facts are ontologically dependent on the values inherent in the holistic practices of inquiry through which we seek to formulate such statements. Here, admittedly, White's primarily epistemological holism is developed into a more robustly ontological holistic pragmatism – but pragmatist approaches to the realism issue in my view more generally integrate ontological and epistemological perspectives (see also Pihlström 2009, 2021).

---

historiography, drawing attention to the practices of historians, is more clearly pragmatist than White's, while Kuukkanen's (2015) postnarrativism seems to be in many ways close to White's position.

[17] Again, this is not far from Kuukkanen's pragmatism. Kuukkanen (2015, 187) also acknowledges that historians seek cognitive warrant for their theses about what is taken to be objectively real, while constructivism is right to insist that "historiographical objects are dependent on the historian's activity". More radical constructivism, such as Roth's (2020), runs the risk of losing the objectively real entirely.

Moreover, when evaluating an argument with a normative conclusion, we may, the holistic pragmatist maintains, revise or reject not only its normative but also (one or many among) its factual premises, provided that relevant adjustments are then made elsewhere in the overall belief system.[18] According to holistic pragmatism, philosophical ideas and arguments also constitute a holistic totality and can only be critically evaluated as elements of such a whole. This holistic epistemology plays an obvious role in the philosophy of historiography, but it is relevant in the other humanities, too. It is – if not according to White's own holistic pragmatism, at least according to more strongly historicist and constructivist pragmatisms such as Joseph Margolis's (e.g., 1995, 2012, 2019) – equally relevant in the natural sciences as well, since there is no inquiry, however objective, that would be immune to the values and interests guiding our purposive practices. Even the natural sciences cannot simply register ready-made objective facts, as pragmatists have always acknowledged. However, the fact that the humanities, including history, cannot do so is not a reason to claim that they would not be in the business of realistically capturing humanly speaking objective truths, or would be "less objective" than the sciences.

Holistic pragmatism thus clarifies not only the epistemology but also the ontology of historiography. A pragmatic realism about historical facts need not, and should not, deny the necessary dependence of the possibility of such facts and truths on the normative perspectives guiding our choice of relevant facts – that is, on the (themselves historically reinter-

[18] A better-known version of pragmatic holism is Quine's (1980 [1953]) narrower view, according to which the unit of empirical testing in science is a scientific theory as a whole, or even the scientific worldview as a totality, instead of any single theoretical statement; the original formulation of this holism is often attributed, also by White, to Pierre Duhem. Quine leaves normative beliefs out of this picture, but for White (2002) irreducibly normative disciplines, such as ethics, history, as well as epistemology itself, are as firmly empirical as science, because (e.g.) our ethical experiences of obligation may be among the empirical "observations" requiring us to adjust our web of belief. This has a tremendous effect on how we view the ethical responsibility of the historian in "choosing" their facts, but such issues cannot be further discussed in this paper.

pretable) value-laden relevance criteria embedded in our practices of inquiry. Holistic pragmatism makes sense of the idea that, when critically evaluating our beliefs and theories (in the humanities or anywhere else), we are always evaluating the entire system of normative principles, including our interests, needs, and values, that structure our inquiries into objective facts. Truth and objectivity are not sacrificed, but their necessary embeddedness in normative interests and values is incorporated in the epistemology of holistic belief fixation and assessment. Furthermore, what White calls historical "memorability" – that is, something's being valuable or relevant enough to be historically remembered – is itself an object of value-laden yet objectivity-pursuing inquiry. There may be no *fully* "objective criteria of historical importance" (in the superhuman sense of "objectivity" demanded by the metaphysical realist), as White says, but there is no reason to suppose that we could not engage in value-guided rational debates on such criteria. A *reflexive* conception of holistic pragmatism presupposes that such pragmatic inquiry into values themselves, as guiding our choices of what is historically relevant and thus real, is possible for us. Hence, we may discuss the rational criteria of historical "memorability" or related relevance considerations as objectively as we can discuss any value-laden issues of the human world.

Let us briefly return to a critical examination of the (historical) antirealism that Orwell's dystopic figure O'Brien espouses. As noted above, Winston claims to remember that airplanes existed already before the Party came to power, while O'Brien denies this, claiming that if the Party says that airplanes have only existed during the reign of the Party, then this is true; no contrast between what the Party claims to be the case and what "really" is the case is possible. A holistically pragmatist analysis can respond to this challenge without either resorting to metaphysical realism (reading Orwell as a straightforward advocate of realism) or being carried away by the Rortyan claim that truth is unimportant in contrast to freedom (Rorty 1989). The claim that the opinion of the Party, or any other contingent authority, partly or entirely constitutes the truth about airplanes is already to apply a certain value-laden framework to this particular issue. Insofar as normative statements about value are in principle critically

tested as elements of one and the same totality together with factual statements, then we are never merely testing the truth-claim about the airplanes (for instance) but an entire conjunction of statements about airplanes *and* the value framework within which it is *possible* for us to claim any truths (or falsehoods) about them.[19] If we take holism seriously, we are inevitably testing our conception of truth itself when testing any truth-claims, as that conception belongs to the same totality of our beliefs facing the "tribunal of experience" as a "corporate body". Only (holistic) pragmatism can, I submit, make sense of this reflexivity as a structural element of our practices of inquiry.

Carried to its logical conclusion, this conception of holistic pragmatism also yields a holistic conception of the relation between what is constitutive of inquiry and what emerges as a result of inquiry. Whenever testing any empirical claim or ontological postulation – in history or any other area – that a certain normative framework of inquiry makes possible for us, we are testing the entire conjunction of the claim or postulation *plus* the framework. Accordingly, our entire conception of pragmatism as a theory of inquiry is itself at stake, holistically, whenever we set out to discuss any particular ontological postulations within our pragmatist account of (say) historiography (or any other field). For the pragmatist philosopher of the humanities, it is fundamentally important to take this reflexive holism very seriously.

Conclusion

This essay has examined a number of interrelated issues in pragmatist philosophy of historiography: the nature of historiographical truth (especially in relation to Kuukkanen's proposal to replace truth by warranted assertion), Peircean realism about "real generals" as an attempt to make sense of the ontology of historiography, as well as the holistic pragmatism formulated by Morton White as a theory of historiographical inquiry dealing with both factual and normative

[19] I have on earlier occasions discussed both Orwell's novel and its relation to realism and truth (e.g., Kivistö and Pihlström 2016, chapter 5) and holistic pragmatism in relation to pragmatist views on truth, values, and normativity (e.g., Pihlström 2021).

matters. These themes have only been briefly and tentatively discussed. A more comprehensive account, embedding the philosophy of historiography in (pragmatist) philosophy of the humanities generally, can be found in my recent book (Pihlström 2022). Moreover, many of the themes only briefly taken up here (or in that book) need further investigation – including the ontological relevance of colligatory concepts in historiography, the exact relationship between the ontological and the epistemological dimensions of holistic pragmatism, as well as the notoriously problematic pragmatist conception of truth in its varieties.[20]

University of Helsinki

References

Ankersmit, Frank (2009). "Narrative and Interpretation." In *A Companion to the Philosophy of History and Historiography*, edited by Tucker, 199–209.

Brandom, Robert ed. (2000). *Rorty and His Critics*. Malden, MA and Oxford: Blackwell.

Colapietro, Vincent (2016). "The Historical Past and the Dramatic Present: Toward a Pragmatic Clarification of Historical Consciousness." *European Journal of Pragmatism and American Philosophy* 8:2, https://journals.openedition.org/ejpap/623.

Conant, James (2000). "The Last Intellectual in Europe: Rorty on Orwell." In *Rorty and His Critics*, edited by Robert B. Brandom. Malden, MA and Oxford: Blackwell.

Dray, William (1980). *Perspectives on History*. London: Routledge & Kegan Paul.

Gronda, Roberto and Viola, Tullio eds. (2016). *Symposium: Pragmatism and the Writing of History. European Journal of Pragmatism and American Philosophy* 8:2, https://journals.openedition.org/ejpap/623.

James, William (1975–78). *Pragmatism: A New Name for Some Old Ways of Thinking* (1907) and *The Meaning of Truth: A Sequel to* Pragmatism (1909). Edited by Frederick H. Burkhardt, Fredson Bowers, and Ignas

[20] Thanks are due to two anonymous reviewers for helpful comments as well as the audience at the Oulu colloquium in June 2022 for interesting discussions.

K. Skrupskelis. Cambridge, MA and London: Harvard University Press. (Single-volume paperback edition 1978.)

Kivistö, Sari and Sami Pihlström (2016). *Kantian Antitheodicy: Philosophical and Literary Varieties*. Basingstoke: Palgrave Macmillan.

Kuukkanen, Jouni-Matti (2015). *Postnarrativist Philosophy of Historiography*. Basingstoke: Palgrave Macmillan.

Kuukkanen, Jouni-Matti (2017). "Moving Deeper into Rational Pragmatism: A Reply to My Reviewers." *Journal of the Philosophy of History* 11, 83–118.

Kuukkanen, Jouni-Matti, Giuseppina D'Oro, Allan Megill, Herman Paul, and Marek Tamm eds. (2019). *Special Issue: Pragmatism and the Philosophy of History*. *Journal of the Philosophy of History,* 13:3.

Margolis, Joseph (1995). *Historied Thought, Constructed World*. Berkeley: University of California Press.

Margolis, Joseph (2012). *Pragmatism Ascendent: A Yard of Narrative, a Touch of Prophecy*. Stanford, CA: Stanford University Press.

Margolis, Joseph (2019). "Pragmatism and Historicity." *Journal of the Philosophy of History* 13, 302–324.

Mounce, H.O. (1997). *The Two Pragmatisms*. London and New York: Routledge.

Orwell, George (1992 [1949]). *Nineteen Eighty-Four.* London: Penguin.

Peirce, Charles S. (1931–58). *The Collected Papers of Charles Sanders Peirce*. 8 volumes. Edited by Charles Hartshorne and Paul Weiss (volumes 1–6) and Arthur W. Burks (volumes 7–8). Cambridge, MA: Harvard University Press.

Peirce, Charles S. (1992–98). *The Essential Peirce*. 2 volumes. The Peirce Edition Project (Nathan Houser *et al*.). Bloomington: Indiana University Press.

Pihlström, Sami (2009). *Pragmatist Metaphysics: An Essay on the Ethical Grounds of Ontology*. London and New York: Continuum.

Pihlström, Sami (2021). *Pragmatist Truth in the Post-Truth Age: Sincerity, Normativity, and Humanism*. Cambridge: Cambridge University Press.

Pihlström, Sami (2022). *Toward a Pragmatist Philosophy of the Humanities*. Albany: SUNY Press.

Quine, W.V. (1980 [1953]). *From a Logical Point of View*. Rev. ed. Cambridge, MA and London: Harvard University Press.

Rorty, Richard (1989). *Contingency, Irony and Solidarity*. Cambridge: Cambridge University Press.

Roth, Paul A. (2020). *The Philosophical Structure of Historical Explanation*. Evanston, IL: Northwestern University Press.

Tamm, Marek (2014). "Truth, Objectivity and Evidence in History Writing." *Journal of the Philosophy of History* 8, 265–290.

Topa, Alessandro (2016). "'I Have To Confess I Cannot Read History So': On the Origins and Development of Peirce's Philosophy of History." *European Journal of Pragmatism and American Philosophy* 8:2, https://journals.openedition.org/ejpap/623.

Tucker, Aviezer ed. (2009). *A Companion to the Philosophy of History and Historiography*. Hoboken: Wiley-Blackwell.

Viola, Tullio (2020). *Peirce on the Uses of History*. Berlin: de Gruyter.

White, Hayden (1973). *Metahistory*. Baltimore: The Johns Hopkins University Press.

White, Hayden (1978). "The Historical Text as a Literary Artifact". In White, *Tropics of Discourse: Essays in Cultural Criticism*. Baltimore, MD: The Johns Hopkins University Press.

White, Morton (2002). *A Philosophy of Culture: The Scope of Holistic Pragmatism*. Princeton, NJ: Princeton University Press.

White, Morton (2005). *From a Philosophical Point of View: Selected Studies*. Princeton, NJ: Princeton University Press.

The Flow of Time and Historical Imagination: Peter Winch on "Ceasing to Exist"

OLLI LAGERSPETZ[1]

1. Introduction

This essay addresses the conceptual preconditions of making a meaningful truth claim about a past event. It mainly considers Peter Winch and his discussion of that question in his lecture, "Ceasing to Exist" (Winch (CTE) 1987). In that text, he is not directly addressing the epistemology of historical knowledge, but the issue he discusses is certainly relevant to the theme of the present volume. My essay aims in part to present Winch's argument and, in part, to connect it with the question of how historical knowledge is possible.

To assert that something has happened is to make a knowledge claim. However, the fact that we can formulate a grammatically correct sentence that looks like a knowledge claim and might, in some situation, express one, does not in itself guarantee that the sentence *is* a meaningful knowledge claim in the situation we are currently facing. The question here is what kinds of background or surroundings we need for something to *count as* an intelligible assertion about what has happened. The main contention in the present paper is that the statement must be related, or at least relatable, to our ideas of how the world generally "works", with recognizable connections with past and future developments. To put this in a different way: A meaningful statement about what has happened never gives us just a "snapshot", but always im-

[1] ORCID https://orcid.org/0000-0002-0317-2268

plies a bigger picture beyond it. The statement must have a place in a flow of time.

R. G. Collingwood was expressing a similar idea when he pointed out that historical research necessarily involves "*a priori* imagination" ([1946] 1989, 240–241). For Collingwood, the crucial aspect of imagination lies in the ability to *interpolate* and *extrapolate*. If we know that Julius Caesar was in Gaul one day and that he was in Rome later, we will conclude that he, at some point, traversed the distance from Gaul to Rome. We *know* he did, despite the fact that we might have no documentation of the trip itself. Similarly, if we see a ship on the horizon and, a few minutes afterwards, see the ship again in a different place, we are "obliged to imagine it as having occupied intermediate positions when we were not looking" (Ibid.). The reverse side is that we cannot seriously assert that Caesar was in two places at once, nor that he was first in Gaul and then in Rome *without* having traveled from the one place to the other. We can say those words, of course, but they would not add up to a meaningful description of a scenario (there is no scenario such as "Caesar being in two places at once"). Historical knowledge implies the assumption of *necessary* connections between states of affairs at different points in time.

The idea of historical *a priori* imagination is relevant with regard to a currently central question in the epistemology of history. Given the fact that the past is not *present*, how is it possible for us to know the past? Won't our knowledge of the past always be *indirect*? We have direct access to documents from the past, but do we have access to the past itself? In a recent paper (forthcoming), Jonas Ahlskog describes this issue as a "shadow of metaphysical realism". While metaphysical realism is probably not the majority position among philosophers of historiography, many of the main questions tend still to be formulated in its terms. The ideal of "direct knowledge" spells out the typical realist contention that the privileged form of knowledge consists in a kind of immediate presence. The ideal form of it would be something like tactile contact or direct eye contact in the present moment. Important early proponents of this view include G. E. Moore (who enlisted it in support of realism) and Bertrand Russell. Knowledge, in its most basic form, would consist in the sim-

ultaneous "compresence" of the epistemic subject and object (see Collingwood [1939] 2002, 25). Collingwood was strongly critical of that view. His idea of historical imagination highlighted the fact that an experience does not constitute knowledge if it is considered outside of context. Our knowledge of any event, however immediate, always incorporates the knowledge of other events, without which it would not make sense to assert it. This means that our experience of the present has no special position above our other sources of knowledge. Making sense of one's own present or recent experiences involves a priori imagination just as much as making sense of the experiences of others in the past.

In his essay on "Ceasing to Exist", Winch does not mention Collingwood's concept of *a priori* historical imagination. Nevertheless, in fact, he presents a detailed analysis of what that concept would mean in a concrete case. Winch considers the imaginary case of a shed that "simply" disappears without a trace. He asks what background we would need in order to understand this scenario – if "scenario" indeed is the right word. His answer is that no amount of direct observation would suffice to make the assertion that things may "simply" disappear so much as intelligible. Our observations, direct or otherwise, are intelligible (and thus, count as "observations" rather than illusions) because they situate the phenomenon in a recognizable time flow between past and future.

This, of course, implies that the epistemological contrast between historical knowledge and knowledge of the present is to some extent a red herring. Both cases of knowledge involve interplay between the present, which we know, and the past, which we also know, as well as an understanding of the general relationship between them. For example, when I gave this lecture in Oulu, my knowledge that I indeed *was* in Oulu was not knowledge of the "present" in the restricted sense that the doctrine of direct knowledge would presuppose. It implied the past experiences of receiving the confirmation that my paper was accepted for the conference, of boarding a train, and so on; as well as a huge amount of knowledge about geography and traveling, plus everyday skills like reading a text and reading the clock. My mere experience of standing in a lecture hall would not have given me the knowledge that I was in Oulu.

Winch asked in his essay what it takes for something to be a meaningful statement of what has happened. He argued that a claim about what has happened situates the event in a broadly "causal" understanding of how the world works. Afterwards, his paper prompted a discussion of what kinds of truth claim his position would rule out. If he referred to a "causal" understanding of the world, would he rule out things like miracles? As his subsequent reply shows, he simply meant that the meaningful description of any event must situate it in *some* intelligible flow of time. As also Collingwood puts it in an unpublished manuscript, a meaningful statement belongs to "a system of thought", in which "every judgment is colored by all the others" (1917, 12).

2. Peter Winch on "Ceasing to Exist"

Peter Winch was famous for his early book *The Idea of a Social Science* (Winch 1958). His comments on the philosophy of historiography in that book are limited to the general question of what it means to understand human agency. Another work by him with relevance to historical knowledge is his late lecture "Ceasing to Exist" (Winch (CTE) 1987). To be sure, this piece does not directly address historical knowledge. However, he engages in a debate with the classical metaphysical tradition, asking how our knowledge of the present relates to the past and the future.

Winch was an important developer of Wittgenstein's philosophy. That philosophical tradition has generally professed an anti-metaphysical stance. Yet, unlike some of his colleagues, Winch had a genuine interest in the metaphysical tradition. He regularly engaged in dialog with the classics, not just criticizing them. It is fair to say that while, like Kant, he rejected ontology as an independent area of study (Winch 1995, 212), he was interested in metaphysics "as natural disposition" (Kant (CPR) 1929, B22). In "Ceasing to Exist", he attempted to describe the conditions of meaningful descriptions of what has happened *without* subscribing to ontological commitments such as naturalism or realism. Unlike Kant, however, Winch did not aspire to a systematic description of the conditions of meaningfulness. On the contrary, he wanted to highlight their great sensitivity to context.

The essay "Ceasing to Exist" raises the question whether we can meaningfully imagine an object purely and simply *ceasing to exist* – just vanishing from one moment to the next. Winch connects his discussion to the magic realism of Nobel-winning author Isaac Bashevis Singer. Singer's short story, "Stories from Behind the Stove", is set in a rural Jewish community in pre-War eastern Poland. The local shopkeeper has a shed in his backyard for garden tools and firewood. It has been part of his life for years, but one day he discovers that it is gone without a trace. Where there once was a heavy log construction, there's suddenly an untouched piece of ground. Singer's story shows us the villagers reacting to this unexpected event – if "event" is the right word.

We might think that the verification of any claim about a present state of affairs simply takes place in the present. In other words, it does not involve a *logical* contradiction to say, on the one hand, (1) "The shed was on the meadow on Monday" and, on the other hand, (2) "The shed was not on the meadow on Tuesday". Moreover, it is not self-contradictory to say that (3) the shed was not dismantled, it did not burn down, etc. on the night between Monday and Tuesday. These three statements do not *formally* contradict each other. Still, it seems to us that the three statements cannot all be true at once.

When Winch juxtaposes the statements, "the shed was there on Monday" and, "the shed was not there on Tuesday", his chief idea is that there is an invisible thread running from the one to the other. We can imagine the transition only by interposing a process where the shed is dismantled, is hit by a bomb, or something else happens – there must be some kind of intervening event.

In other words, a factual statement that seemingly just concerns a single observation in the present implies, in practice, unstated assumptions about what kinds of thing sheds are and what meadows are. Generally, it implies a world as a place held together by a mesh of causal relations, including the identities of individual objects persisting in time. This is why the sentence, "The shed has simply vanished" would imply a challenge to a complete world-picture. Our understanding of the world implies *necessary* relations between the past and the present. If, where we expected to find a shed,

there is just a meadow with no traces of a building, the implication is that there has not been a building there at all for a long time. Indeed, just to say that something is a *meadow* is to presuppose ideas about the natural growth of plants, the approximate time it takes, and the like. Anyone would agree to these simple conceptual points unless they have some hopeless philosophical ax to grind. At the same time (even if Winch does not say so), they testify to the existence of metaphysics, in our everyday lives, as a "natural disposition". We just *cannot* agree that descriptions of the present have nothing to do with descriptions of the past and of the future.

In his lecture, Winch employs the expression "the stream of life". He quotes Wittgenstein on the concept of verification:

> The stream of life, or the stream of the world, flows on and our propositions are so to speak verified only at instants. Our propositions are only verified by the present. So they must be so construed that they can be verified by it (Wittgenstein (PR) 1974, V: § 48; quoted in Winch, (CTE) 1987, 98; Winch 1989, 22).

To verify a proposition is to dip one's toe in the flow of the world. Any momentary observation would not have the same implications – it would in fact *not be* the same observation – if the flow was different or absent. Famously, a central idea in Wittgenstein's philosophy of language was that a sentence could only have meaning in a context. A string of words, considered in the abstract, does not reveal us how, or whether, it constitutes an assertion about reality. Quite analogously, a single experience, considered on its own and taken out of the general flow of life, does not produce knowledge. Verification "at an instant" must have a connection with some kind of possible inquiry where it would solve a question.

A heavy wooden structure does not burst like a soap bubble. *Something* must be an illusion: either the villagers' perception that the shed does not exist now, *or* their previous memory of it as existing – or both. To be sure, perception takes place in the present, but it becomes *verification* only through its connections to intelligible causal relations stretching from the past to the future. In a certain sense, *no perception is confined to the present*.

These observations are at odds with the doctrine of immediate knowledge mentioned here in the Introduction. Witt-

genstein's idea was that we *identify* our perception of the present against the background of what else we know to be relevant in the general situation. The doctrine of immediate knowledge holds, on the contrary, that we *build up* our general idea of reality out of immediate observations. These "pure" cases are as it were the Lego bricks of knowledge. Reality enters the mind of the subject *directly* and provides the subject with secure grounds for further conclusions. As Michael Dummett has put it, these secure grounds, at "the most elementary level", can be expressed as basic "observation-sentences": "our original grasp of there being something that makes a statement true derives from our use of basic forms of statement as reports of observations" (Dummett 1973, 465–467; quoted in Winch 1987, 41).

A central question here is in what way we can say that direct acquaintance with the present provides us with *knowledge*. In Wittgenstein's own time, Moore and Russell argued, each in their own way, that the ultimate source of knowledge consisted in acquaintance with objects that directly confronted us (Moore [1910–11] 1957, 122–123; Russell 1905, 492; 1923). Direct knowledge supposedly just "is" there: it relies neither on a reasoned assessment of sources nor on weighing different theories against each other. Wittgenstein pointed out that this assimilates knowing to a kind of seeing, where a "*fact*" is directly "taken into my consciousness" (Wittgenstein (OC) 1969, § 90). For Wittgenstein, as for Collingwood, knowledge was, instead, the product of practices of inquiry, such as assessing sources, their implications, and their reliability (see Lagerspetz 2021, 56). In this way, knowledge is not neatly separable from the ability to give (at least possible) grounds. From the point of view of Wittgenstein and Collingwood, it was more plausible to think of direct intuition as not knowledge at all. At most, it was as a problematic, *limiting case* of knowledge.

3. The Role of Memory

At this juncture, Winch cites Kant's "Second Analogy", included in his *Critique of Pure Reason* (Kant (CPR) 1929,

B232/A189–B256/A211; Winch (CTE) 1987, 100).[2] A memory episode is not merely a sequence of mental images. Remembering a series of events is also to relate it meaningfully to a shared, intersubjective world. This is why Kant argues that understanding the sequence of images as a sequence of *memories* implies the concept of causation. Things must hang together in an intelligible natural order in order for the sequence to represent a chain of events.

As an example, Winch describes his memories of a trip from Manhattan to London. Let us imagine, he says, that he remembers his trip in the following order. *First*, he takes *the subway* from Manhattan to central London, *from there* a flight to the JFK in New York, *and from there again* he goes by *Tube* to Earl's Court station in London (Winch (CTE) 1987, 101). It is quite possible that the mental images come to his mind in exactly this order. Yet, in the light of geographical facts, he would immediately correct his memories to agree with facts. "[T]he impression", he says, "however overwhelmingly strong, that *this* is what happened by no means has final authority"; otherwise, we would not be able to "distinguish such a narration from a fantasy" (Winch (CTE) 1987, 102).

This, of course, calls for some clarification. Admittedly, it is possible to produce a film or a cartoon showing Winch traveling in exactly this order. Similarly, the cartoon might show a building and, in the next frame, show the same site without the building. In that sense, we can picture the disappearance of a shed to ourselves (i.e., we can imagine two distinct situations and include them in a sequence of pictures). However, Winch argues that the succession of two images does not represent an *event* of any kind.

Similarly, we can string together a sentence, and, in that limited sense, we can "say", for instance, "The shed simply vanished". Winch is asking whether we could use the sentence to represent *an assertion*. In other words, do we under-

---

[2] See especially B234: "In other words, the *objective relation* of appearances that follow upon one another [in my perception of a temporal sequence] is not to be determined through mere perception" ... Experience itself … is thus possible only in so far as we subject the succession of appearances, and therefore all alteration, to the law of causality". See also Winch 1989, 52.

stand what it would mean for someone seriously to state that "this" (whatever it is) has happened? As Winch puts it elsewhere,

> Of course, I can perfectly well arbitrarily utter certain words which, uttered in other circumstances, would constitute an assertion. But to the same extent as you thought I *had* uttered them arbitrarily, you would be disinclined to think that I *had* made any assertion (Winch 1987, 40).

In "Ceasing to Exist", Winch argues that the sentence, "The shed vanished", does not represent an event, but merely one's bafflement in the face of what one cannot make sense of. He describes it like this:

> If we look again, more closely, at Singer's story we shall see that we are really offered no more than the title "the shed vanished" to the story Zalman [the narrator] is depicted as spinning (Winch (CTE) 1987, 91).

The owner, coming to inspect the site, does not start by concluding, "the shed has vanished" but instead, "I must have lost my mind" (Winch (CTE) 1987, 92). All that remains is the observation, "We don't understand what has happened". The words, "It has vanished", are a way to voice this utter bewilderment.

4. Miracles and Unheard-of Occurrences

"Ceasing to Exist" gave rise to a debate that engaged, in particular, colleagues in the Wittgenstein tradition where Winch was mainly working. The debate included a clarifying response by Winch (Holland 1989; Holland 1990; Malcolm 1990; Marshall 1990; Mounce 1988; Mounce, unpublished; Palmer 1995; Phillips 1993; Phillips 1993b; Winch 1995). The debate focused on the philosophy of religion, especially on the concept of a miracle. This had not been at the center of the original lecture, even though Winch did comment on it, making references to R. F. Holland's (1980) essay on "The Miraculous" (Winch (CTE) 1987, 94, 95, 96).

At least according to one quite usual definition, a miracle would be an unexplained occurrence in defiance of the laws of nature. In other words, what *did* Winch say? Did he say

that a miracle not only eludes any natural explanation, but that a miracle cannot even be coherently described or even imagined?

Winch (1995) clarified his position in his contribution to the debate. His idea was that the intelligible description of any event would place that event in some relation to a general worldview. For instance, we have an idea of what kinds of structures sheds are and what one can expect to happen to them. A shed that radically departs from expectations is not a shed, but something else. It might be an illusion of some kind. We adjust all our statements about reality to whatever *can* constitute part of 'the stream of life' or 'the stream of the world' as we understand it.

In a modern industrial society, this would, at least often and generally, mean that we rule out supernatural occurrences: our world must conform to physical "laws of nature" (even if this kind of naturalism is probably less pervasive in our lives than we tend to think it is). In the village of Singer's short story, the belief in demons was widespread, and religion had an influence on thinking. In a religious community, the concept of a miracle is a legitimate one. A miracle is, however, not the same as an unexplained or unintelligible occurrence, for it does have an explanation. *God* performs the miracle, perhaps through a representative. Religious traditions presuppose that events reported as miracles have some religious significance. Singer's villagers did not think that they had witnessed a miracle because, as they saw it, God would not conceivably take an interest in a random object like a shed.

Useful applications of the concept of a miracle also presuppose a culture where there is room for informed discussion of would-be miracles. The impossibility, for many of us inhabiting industrial societies, of accepting *any* account of a miracle testifies to our reluctance to attribute *anything* – miracle or no miracle – to Divine agency.

5. Correspondence or Coherence?

The ensuing debate on the philosophy of religion did not really touch on the most important question Winch was raising. He did not propose a theory of what kinds of occurrence are

possible, but rather he was asking what background we need if we are to claim that something has happened. Putting this in more general terms: what is involved in the informed statement that something is the case?

One answer might look correct in its rough outlines. Claims about reality aim to "correspond with" reality. There must be *something* in reality with which we should compare the claims we make. When we think, we think of an object. Our thought is correct if it corresponds with the object. For Winch, that answer was correct *in a sense* and indeed a truism: "we cannot simply assert anything at will" (Winch 1987, 40). We might think this point implies support of the *correspondence theory* of truth. However, Winch points out (also in his other work) that the idea of correspondence tends to paper over important complications:

> It is one thing for a man to think that something is so and quite another thing for what he thinks to *be* so. This simple truism is fundamental to what we understand thought to be; for a thought is a thought about something – it has an object – and the kind of relation it has to its object involves the possibility of *confronting* it with its object. […] However, it is considerably easier to recognize this as a truism than it is to understand exactly how it is to be applied in different areas of human thinking. The attempt to win clarity about such issues is *philosophy* (Winch 1987, 194, italics added).

The correspondence theory of truth at best lays out a research question – the question of what it *means*, in various cases, for our ideas to make contact with reality. The theory itself, in this general form, does not provide the answer. At the same time, the theory muddies the waters by introducing the blanket idea of correspondence, as if one could reduce complex relations between thinking and reality to just *one* kind of relation. To "compare thinking with reality" amounts to very different things in different situations. We do not make progress without at first asking what "comparison" would mean in the case we are facing.

Note, however, that the argument in "Ceasing to Exist" hits equally at *coherence* theories of truth. In a narrow sense, "The shed was on the meadow on Monday" is coherent with "The shed was not on the meadow on Tuesday" as well as

with "Nothing happened to the shed between Monday and Tuesday". There is no *formal* contradiction between the three sentences. However, there *is* a kind of contradiction at play because we understand that we cannot coherently assert the three statements at once. But once more, coherence theory just states the problem but does not give the answer. The question here would be why we think of the three sentences as mutually incoherent, despite the fact that there is no formal contradiction at play. The crucial thing is that we presuppose the background of something like an ordered universe. Neither correspondence nor coherence gives us *that*.

In "Ceasing to Exist", with these remarks on the relation between experience and the flow of time, Winch in fact continued a discussion he already started in his first book, *The Idea of a Social Science*. In that book, he described the main task of philosophy as that of clarifying "man's [sic] relation to reality", in other words, "the *force of the concept* of reality" (Winch 1958, 9). He pointed out already in that book that science, religion, art, etc. all aim at clarifying our relation to reality, without that implying that they all are branches of a single enterprise. The philosophically interesting task is to discern the various forms that "making contact with reality" may take. In "Ceasing to Exist", he suggested that even the activity that we usually take to be unproblematic – direct observation – holds no privileged position. The significance of direct observation is dependent on how it connects with our various ways of making sense of the world.

5. Conclusion

The relevance of this discussion to the epistemology of history lies in the light it sheds on the idea of experiencing the present. Due to the influence of metaphysical realism, many philosophers of history feel the need to justify historical knowledge in realist terms. The guiding assumption is that the superior form of knowledge is one where the object of knowledge is "available" directly in front of us in the here and now. Through this privileging of the present, a kind of skeptical challenge is introduced.

From that point of view, it would appear that historical knowledge is *direct* only in its relation to records currently in

existence: artifacts, written sources, etc. – and apparently only at the very moment we are studying them. It would not be so in relation to the events themselves. But then, as Ahlskog puts it in a critical discussion, "how can this present experience ever get in touch with the past that it is supposed to uncover?" (Ahlskog 2021, 103). Scholarship would give us an interpretation and a narrative, but, supposedly, "context kills authenticity" (Ankersmit 2005, 172, 180; quoted in Ahlskog 2021, 101). One alternative, recently offered by *presence theory*, would be that historians should try to recover a "presence of the past". This would be a direct relationship with the past "predicated on our unmediated access to actual things that we can feel and touch and that bring us into contact with the past". Presence theory is framed as "an attempt to reconnect 'meaning' with something 'real'" (Kleinberg 2013, 11; quoted in Ahlskog 2021, 100). Again, the implication is that only what is present is real.

These theories are based on the insistence that our knowledge of the past is inherently problematic, while our knowledge of the present is not. However, the problem is shown in a different light once we see that our knowledge of the present has pretty much the same complexities as our knowledge of the past. The skeptical question was whether any knowledge can exist, other than knowledge of the here and now. The implicit answer that we can get from Winch is that there *is* no such thing as knowledge of the here and now; that is, not in the free-floating sense that the question would presuppose. Our knowledge "of the here and now" always presupposes some knowledge of the past and the future, and an idea of their connections. There are important parallels between Winch and Collingwood on this issue, especially – as I have noted – considering Collingwood's critique of realist theory of knowledge and his concept of historical imagination.

Acknowledgement

Part of this work was written during my research stay at the Center of Ethics, the University of Pardubice, Czech Republic, in 2022. The project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No.

101026669 (WC-Cult). Thanks to Weston Library, Oxford, and to Teresa Smith for use of the Collingwood manuscripts.

Åbo Akademi University

References

Ahlskog, Jonas (2021). *The Primacy of Method in Historical Research: Philosophy of History and the Perspective of Meaning*. Routledge: New York.

Ahlskog, Jonas (forthcoming). "How History Is about Reality." *Journal of the Philosophy of History*.

Ankersmit, Frank (2005). *Sublime Historical Experience*. Stanford: Stanford University Press.

Collingwood, Robin George (1917 unpublished). "Logic and Contradiction." Weston Library, Oxford University, Dep. Collingwood 16/1.

Collingwood, Robin George (1939/2002). *An Autobiography*, Oxford: Clarendon Press.

Collingwood, Robin George (1946/1989). *The Idea of History*, Oxford: Oxford University Press.

Dummett, Michael (1973). *Frege, Philosophy of Language*. London: Duckworth.

Holland, Roy Fraser (1980). "The Miraculous." In *Against Empiricism*, edited by Holland, R.F. Totowa, NJ: Barnes & Noble Books, 169–187.

Holland, Roy Fraser (1989). "Lusus Naturae." In *Wittgenstein: Attention to Particulars*, edited by Phillips, D. Z. & Winch, P. London: Macmillan, 45–60.

Holland, Roy Fraser (1990). "Naturalism and Preternatural Change." In *Value and Understanding: Essays in Honour of Peter Winch*, edited by Gaita, R. London: Routledge, 32–41.

Kant, Immanuel (1929). *Immanuel Kant's Critique of Pure Reason.* Translated by Smith, N. K. London: Macmillan. Abbreviated (CPR).

Kleinberg, Ethan (2013). "Presence in *Absentia*." In *Presence: Philosophy, History, and Cultural Theory for the Twenty-First Century*, edited by Ghosh, R. & Kleinberg, E., Ithaca, NY: Cornell University Press, 8–26.

Lagerspetz, Olli (2021). "The Linguistic Idealism Question: Wittgenstein's Method and his Rejection of Realism." *Wittgenstein-Studien*, vol. 12, no. 1, 2021, 37–60. https://doi.org/10.1515/witt-2021-0003

Malcolm, Norman (1990). "On 'Ceasing to Exist'." In *Value and Understanding: Essays in Honour of Peter Winch*, edited by Gaita, R. London: Routledge, 1–12.

Marshall, Graeme (1990). "Intelligibility and the Imagination." In *Value and Understanding: Essays in Honour of Peter Winch*, edited by Gaita, R. London: Routledge, 13–31.

Moore, George Edward (1910–11/1957). *Some Main Problems of Philosophy*. Leicester: The Blackfriars Press.

Mounce, Howard O. (1988). "Critical Notice of Peter Winch's *Trying to Make Sense." Philosophical Investigations* 11 (1988), 238–239.

Mounce, Howard O. (unpublished). "On Simply Vanishing." Discussion paper, quoted (by Phillips 1993) "with the author's permission".

Palmer, Anthony J. (1995). "Violations of Nature and Conditions of Sense." In *Philosophy and the Grammar of Religious Belief* , edited by Tessin, T. & von der Ruhr, M. New York: St Martin's Press, 186–199.

Phillips, D.Z. (1993). "Waiting for the Vanishing Shed." In *Wittgenstein and Religion*, edited by Phillips, D. Z. New York: St Martin's Press, 171–192.

Phillips, D.Z. (1993b). "Miracles and Open-door Epistemology." *The Scottish Journal of Religious Studies* XIV (1993), 33–40.

Russell, Bertrand (1905). "On Denoting." *Mind, N.S.* Vol. 14(1905), 479–493.

Russell, Bertrand (1923). *Problems of Philosophy*. London: Williams and Norgate.

Winch, Peter (1958). *The Idea of a Social Science and Its Relation to Philosophy*. London: Routledge & Kegan Paul.

Winch, P. (1987). *Trying to Make Sense.* Oxford: Blackwell. (Includes the chapters: "Im Anfang war die Tat", 33–54; (CTE) "Ceasing to Exist", 81–106; "Language, Belief and Relativism", 194–207.)

Winch, Peter (1989). *Simone Weil: "The Just Balance"*, Cambridge: Cambridge University Press.

Winch, Peter (1995). "Asking too Many Questions." In *Philosophy and the Grammar of Religious Belief*, edited by Tessin, T. & von der Ruhr, M. New York: St Martin's Press, 200–214.

Wittgenstein, Ludwig (1974). *Philosophical Remarks*. Edited by Rhees, R., Oxford: Basil Blackwell. Abbreviated (PR)

Wittgenstein, Ludwig (1969). *On Certainty*. Edited by Anscombe, G.E.M. & von Wright, G. H., London: Basil BlackwelL. Abbreviated (OC).

Still or Sparkling? Past, Present and Future in Bohm's Implicate Order Approach

PAAVO PYLKKÄNEN

> Time is Nature's way to keep
> everything from happening all at once.[1]

1. Introduction

David Bohm (1917–1992) was a physicist who published extensively on philosophical topics. He is particularly well-known for proposing, after discussions with Einstein, the first consistent ontological interpretation of quantum theory (Bohm 1952 a & b; Bohm and Hiley 1987, 1993; Goldstein 2021).[2] But he was also trying to develop a much more ambitious scheme, a new framework or scientific world view in which one could hope to bring together quantum theory and (general) relativity, and include biological and psychological phenomena (Bohm 1980; Pylkkänen 2007). This new "impli-

[1] Discovered among graffiti in the men's room of the Pecan Street Café, Austin, Texas. Reported by John Wheeler (1990), 315.

[2] For a recent biography of Bohm, see Freire, O., Jr. (2019, vi), who writes: "The legacy of …. Bohm to physics may be stated in a nutshell: He was a physicist who made many and lasting contributions particularly in subjects such as plasma, metals, and quantum mechanics; he was one of the discoverers of the Aharonov–Bohm effect and suggested alternative interpretations of quantum mechanics. He was undoubtedly one of the major twentieth-century physicists. […] Bohm's main legacy in quantum mechanics, I think, was his contribution to keep [the interpretation debates in quantum mechanics] alive in times when many of physicists thought it should be closed." For a discussion of Bohm's 1952 interpretation of quantum theory and its relation to relativity, see Goldstein (2021), and Bohm and Hiley (1993).

cate order" world view, combining process philosophy with structuralism, also implies a certain physics-based way to think about time, including how to understand past, present, and future; being and becoming; and the nature of movement. Philosophical discussions of time sometimes take a one-sided view from physics, for example by emphasizing only the special theory of relativity. But a more accurate notion of time in physics should surely try to do justice to all fundamental theories in physics, especially quantum theory and (general) relativity. Bohm was one of those physicists in the latter half of the 20th century who truly struggled to bring out the philosophical (and especially ontological) implications of both quantum theory and relativity and attempted to describe them in a more general way. In this short article I will present and discuss some of Bohm's ideas about time, hopefully helping philosophers of time to evaluate whether they could be helpful in tackling some of the difficult problems connected to our ideas about time (see also Limnell 2008).

2. Quantum mechanics challenges the block universe view and suggests extended, overlapping moments as the spatio-temporal building blocks of the universe

Bohm's implicate order scheme promises to give us the notion of the present moment as genuinely real and extended; the notion of the past as "enfolded" structures in the present moment (nested memories or reverberations or active transformations); and the notion of future as anticipated, "enfolded" potentialities that are actualized in a creative way in the present moment. This is a tall order. The first objection might come from the supporters of the block universe theory. Does not relativity theory say that space-time is a block, and time (in the sense of past-present-future, as well as genuine becoming) is an illusion? Bohm (1986: 183) argues that quantum mechanics strongly questions the foundations of the block universe view. First of all, he claims that the uncertainty principle of quantum mechanics implies that the point event of relativity theory (a key element of the block universe view) cannot in general have an unambiguous meaning. Even more radically, he suggests that Einstein's well-known photon-in-a-

box experiment (when extended) implies a fundamental breakdown in the notion that the point event can be absolute.[3] Instead, Bohm says, quantum mechanics suggests that there exist objective but context-dependent moments which are extended in some sense and ambiguously located in space and time. The term "moment" should here be understood in the same sense as our actual experience of the moment "now", as something that is never completely localizable in relation to other moments, while typically overlapping with them.[4] In the quantum mechanical domain, the extension and

---

[3] An anonymous referee was concerned about the conflict between Bohm's view and the principles of relativity theory. I will not go into this question in detail in this paper, but I include a quote from Bohm's 1986 article, which provides more explanation. Note in particular that in Einstein's photon-in-a-box experiment the breakdown in the notion that the point event can be absolute is a consequence of the general theory of relativity (when we consider quantum indeterminacy in its light):

[A]ccording to modern physics, microprocesses are […] very fast, irregular, and ambiguously related to what comes next. Indeed, it is not in general possible to relate the specifiable information content unambiguously to succeeding events (this is just the essential meaning of the Heisenberg uncertainty relations, as interpreted by Bohr). Here, too, the relevance of the usual notions of time may be questioned. What seems to be called for is that we recognize that the "point event" of relativity theory cannot in general have an unambiguous meaning.

General relativity leads to this conclusion in an even more forceful way. There is Einstein's well-known hypothetical experiment of weighing a photon in a box. To make what is meant more vivid, imagine a box within which is a whole context of process determining a space-time order and measure appropriate to this context. This box is supported on a spring, which has a certain "quantum indeterminacy" or ambiguity in its height above the ground. It is then a consequence of the general theory of relativity that there is a minimum ambiguity or "uncertainty" in the relationship between the rates of processes inside the box and those outside, which latter are based on the more solid support of a firm foundation that does not move.

More generally, this relationship depends on the whole context. This constitutes a fundamental breakdown in the notion that the point event can be absolute." (Bohm 1986: 183)

[4] For a thorough discussion of theories involving overlapping moments of temporal consciousness, see Dainton 2022. By suggesting that overlapping moments are the spatio-temporal building blocks of the physical

duration of a moment is dependent on the quantum mechanical wave function. Each moment is subject to a certain lack of precise localizability over a region in which the wave function is appreciable. (Ibid., 184.)[5]

This idea of moments in the physical universe is similar to our experienced moment of "now", which is also extended

---

world, Bohm (like Whitehead) opens up a new way to understand the relationship between conscious experience and the physical world.

[5] An anonymous referee was concerned about whether causal structure (essential for relativity) might be violated in Bohm's view which includes extended and ambiguously located, overlapping moments. The way that Bohm deals with this is to suggest that ordinary space-time (and ordinary causality) emerge as a limiting case or approximation from the implicate order (or from what John Wheeler calls "pre-space" or "pre-geometry"). Quantizing gravity implies that for very short times and distances, the entire notion of space-time becomes totally undefined and ambiguous. Bohm (1986: 192) notes that Wheeler (1980) in particular has drawn attention to "… a fundamental ambiguity in the meanings of the terms that define not only the rates of clocks and the lengths of rulers, but also relationships of "before" and "after."" Bohm further notes that "…such an inability to define "before" and "after" (along with the light cone itself) dissolves the entire conceptual basis on which our notions of space and time depend." (1986: 192) Or as Wheeler (1990: 315) himself puts it: "What are we to say about that weld of space and time into spacetime which Einstein gave us in his 1915 and still standard classical geometrodynamics? On this geometry quantum theory, we know, imposes fluctuations […]. Moreover, the predicted fluctuations grow so great at distances of the order of the Planck length that in that domain they put into question the connectivity of space and deprive the very concepts of "before" and "after" of all meaning." Wheeler's solution was to characterize space-time as a kind of very fine foam out of which continuous space, time and matter emerge as approximations on the large-sale level. However, because the structure of the foam is given by quantum laws, Bohm thinks it is more accurate to characterize pre-space as a form of the implicate order: "My attitude is that the mathematics of the quantum theory deals primarily with the structure of the implicate pre-space and with how an explicate order of space and time emerges from it, rather than with movements of physical entities, such as particles and fields." (1986: 192-3) In Bohm's view, moments are projected from the implicate order, but they are typically projected in such a way that no violations of causality are possible.

and ambiguously located in space and time and internally related to other moments (e.g., overlapping with them, or enfolding information about them as memories). Bohm implies that we can assume that the present moment, with its extended spatio-temporal structure, is genuinely real. The past of a given present moment can be enfolded as a nested structure (e.g., a memory) that is present in that actual moment. Moments are thus not isolated or merely externally related point events, as in relativity theory, but are instead internally related extended structures and processes.

So let us assume, for the sake of the argument, that the block universe view is mistaken, and that there exist, as Bohm suggests, extended moments, ambiguously located in space and time, overlapping with each other. How do we understand the past, the present, and the future in this way of thinking? And what is this "enfoldment" or "implicate order" that Bohm is relying upon?

3. Time is Nature's way to keep everything from happening all at once: How does Bohm's glycerine tank do the job?

We always find ourselves in the present moment. But when this present moment ends, where does it go? And where does the next present moment come from? Somehow, we need to a have a present moment, but we also need to get rid of it after a while, to make room for the next present moment that is to come. How can we understand this? Can Bohm's implicate order scheme help here?

There is an analogy or a model which helps to illustrate how time can be understood in terms of the implicate order. For Bohm, time is an abstraction from process – time has to be abstracted from an ordered sequence of changes in an actual physical process (Bohm 1986, 189). We are now going to describe a physical process which can be used to illustrate past, present, and future; it also provides a way to think about the present moment, what happens to it when it is gone, and how the new present moment comes into being. The model also illustrates how movement can be understood in a new way; and later in the article we will even briefly show how the model can be used to describe particle creation and annihila-

tion in quantum field theory without having to assume (as the usual view does) that particles are moving backwards in time. But please note that what we are going to describe is merely a model or analogy. It has some mechanical features in it which, for one thing, do not capture some important non-mechanistic features of the quantum processes. So please take the analogy with a Wittgensteinian grain of salt – as a scaffolding or ladder which is to be kicked away once it has helped you to understand the point it is trying to make.

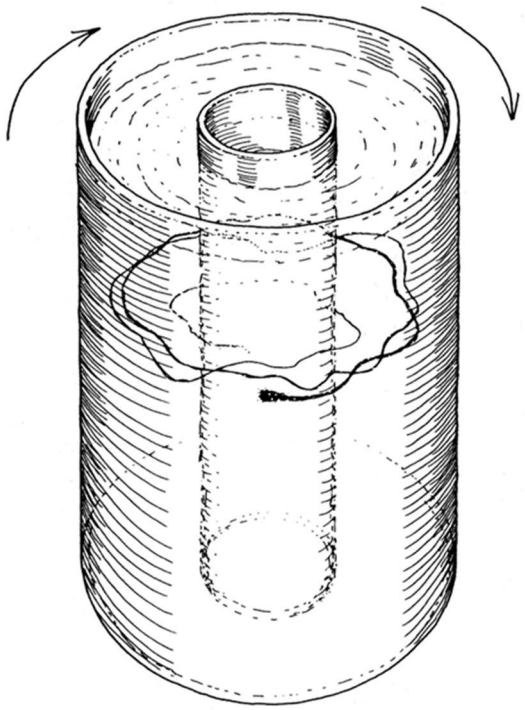

Figure 1. The ink-in-glycerin device

The analogy makes use of an apparently simple device: a tank made of two concentric glass cylinders filled with viscous fluid such as glycerine between them (see Figure 1).[6] The out-

[6] For a link to a *New Scientist* video of this device, see https://transitionconsciousness.wordpress.com/2015/09/19/the-experiment-which-inspired-david-bohm/ In this particular model the

er cylinder can be turned very slowly, which results in negligible diffusion of the fluid. Let us place a droplet of insoluble ink in the fluid and turn the outer cylinder. As we keep turning, the droplet is drawn out into a fine thread-like form that eventually becomes invisible. However, when we turn the outer cylinder in the opposite direction, the thread-like form draws back and suddenly becomes visible, as a droplet essentially the same as the original one (Bohm 1980, 179).

How can we use this device to model past, present, and future? Let us say that when we put the droplet to the glycerine, and do not turn it, it represents the present moment, the state of the universe at the present time t. When we then turn the outer cylinder, the droplet is drawn out until it becomes invisible. It still exists in the present moment, but not as a visible entity. Bohm would say that the droplet has become "enfolded" into the whole glycerine, analogously to the way an egg is folded into the dough when we are preparing a cake.[7] If we think this as a model or time, we can say that what was previously the present moment has been enfolded. It exists in some sense in the present moment, but not in the same way as the original droplet. The original droplet, before turning, was in an unfolded or "explicate" state, while after turning, the droplet is in an enfolded or "implicate" state. Now, let us say that the present moment corresponds to the explicate, and the past to the implicate.

Can we use this device to model the future? To get started with this, let us make a very simplifying assumption that we know what the future will be like. To make a model of the past-present-future, I will start with the device in a state

---

inner (rather than the outer) cylinder is turned. Bohm first saw this device when watching a BBC television programme and made extensive use of it in his 1980 book *Wholeness and the Implicate Order*.

[7] We are dealing with an enfolded or "implicate" order here. Note also that this is a very special kind of enfolded order because we can unfold it back to where we started from, because of the reversibility made possible by the properties of glycerine. In contrast, once we have folded the egg to the dough, we will not get it back by reversing the motion of the whisker, because in that situation the egg has undergone irreversible diffusive mixing. There is a sense in which the egg in the dough still exists as an enfolded structure, but for all practical purposes we have lost the possibility of unfolding it.

where no droplets have yet been placed. First, I will put in the glycerine a droplet which represents the state of the universe at the most distant future I want to model. I then turn n times until the droplet is enfolded. I then put in a droplet which represents the state of the universe at a slightly earlier time in the future and turn n times (which means that the droplet I put in first has been enfolded 2n times). I keep on doing this until a great number of droplets have been enfolded. I finally put in the last droplet, and do not turn. Let us say that this last droplet, which is unfolded and visible, represents the present moment. Now I can start turning the outer cylinder in the opposite direction. The droplet that represents the present moment begins to be drawn out into a fine thread-like form; it will be enfolded and becomes the past. But as I keep turning, the last of the droplets I put in and used to model the future becomes visible first as a thread-like form which draws back and then becomes visible as a droplet. This droplet, which was previously enfolded and represented a future moment, has now become unfolded and represents the present moment. But if I keep turning, this droplet will in turn re-enfold and become a representation of the past, making room for the next droplet which will unfold and represent the present.

So here we have a very simple way of thinking about the past, the present, and the future. The present moment is represented by the droplet that is currently unfolded and visible; the past moments are represented by droplets that have been enfolded n, 2n, 3n etc times. The future moments are likewise represented by droplets that have been enfolded n, 2n, 3n etc. times.

There is a sense in which the past, the present, and the future all exist. But they do not exist in the same way. As we mentioned, the model gives a very simplified picture. For example, the way we have used it implies that the future is determined, existing as well-defined enfolded structures, which will unfold mechanically into successive present moments. Bohm actually thought that the future exists as potentialities, and that the way these potentialities unfold or become explicate in the present can involve genuine creativity. So, it is clear that the scheme has to be developed to allow for this (see Bohm 1986; for a more technical discussion, see

Bohm 1987; for a recent discussion of Bohm's notion of potentiality in the context of theology, see Korpela 2022).

4. Movement understood in terms of the implicate order

We can use this device also to model movement. To do this, let us insert a droplet, A, in a certain position and turn the cylinder n times. Then we insert a droplet, B, in a slightly different position and turn the cylinder n more times (so that A has been enfolded by 2n turns). And then we insert C along the line AB and turn the cylinder n more times, so that A has been enfolded by 3n turns, B by 2n turns, and C by n turns. We proceed in this way to enfold a large number of droplets. Then we move the cylinder fairly rapidly in the reverse direction. If the rate of emergence of droplets is faster than the minimum time of resolution of the human eye, what we will see is apparently a particle moving continuously and crossing the space. But in this analogy, there is no such single particle. What we have is a set of co-present elements ("droplets") at different degrees of enfoldment, giving rise to the appearance of a particle moving continuously. This is a key characteristic and assumption of the implicate order theory: movement typically involves a set of co-present elements at different degrees of enfoldment.

Bohm thinks, radically, that moments of conscious experience and moments of physical reality are analogous to each other. We can experience movement in conscious experience, and we also assume that there is movement in the physical world, taking place independently of the human mind. He suggests that in both cases, the nature of movement is similar in some important respects. What does he mean by that?

As an example, he discusses our experience when we are watching a motion picture, say a short part of a film where a car is moving. The input consists of a set of discrete images which are typically slightly different from each other. If the time interval between the images is short enough, we will experience movement in our conscious experience (i.e., we see a car moving). How does this come about? We might assume that the distinct images are processed in the brain, enfolded as it were, but yet retain their identity in some sense.

Somehow all the images, although different, are sensed together, and this sensing, Bohm suggests, constitutes or gives rise to our conscious experience of movement. As we saw with the glycerine tank, a key feature of an implicate order is that there are "co-present elements at different degrees of enfoldment". The image that is first perceived "now" has the lowest degree of enfoldment (it is most explicate, cf. Husserl's primary impression), while the images that were perceived earlier have been slightly enfolded, and thus have a higher degree of enfoldment (cf. Husserl's retention). We typically also anticipate the images to come; perhaps such anticipations can be understood as implicit, enfolded, not fully explicate structures (cf. Husserl's protention). Thus, we can understand our experience of movement in consciousness by assuming that there are distinct neural representations in the brain, at different stages of processing (or "enfoldment") and that when these are sensed together, a sense of movement arises in conscious experience.[8,9]

---

[8] Another example of the implicate order in conscious experience is listening to music. We hear certain notes for the first time in a given moment. Some of the notes that we have already heard earlier are present in our experience as active transformations rather than memories. Typically, their degree of enfoldment is greater the more time has passed but they are still present. The future is present in the experience as anticipated notes. Just as with the example of the motion picture, sensing the neural representations of the co-present tones at different degrees of enfoldment gives rise to subtle sense of movement (e.g., the movement of a symphony) Bohm (1980). Pylkkänen (2007, ch. 5) has applied the notion of implicate order to describe time consciousness and related it to e.g., Husserl's views (see also Hautamäki 2021).

[9] It would also be interesting to compare Bohm's view of movement in relation to the currently prominent approach to modelling perception, namely predictive processing (see Clark 2013). This approach assumes that perception is not merely a passive process in which information about the environment is received and processed until it gets experienced. Rather, what is important is that the brain is able to apply a hierarchical generative model which enables it to predict the unfolding (incoming) sensory stream. Presumably in this approach our experience of movement is assumed to be constituted more of the properties of the predictive model, than from the elements in the unfolding sensory stream. Whether the implicate order could be used to explain how the experience of movement

How is the above analogous to movement in the physical world? Think again about the glycerin tank, and the case where we placed a droplet in a certain location and turned the outer cylinder n times to enfold it, then a second droplet in a very nearby location, enfolding it in a same way, and then a third one along the line defined by the two previous ones etc. When we turn back quickly enough, it appears to us as if there is a "particle" moving in a straight line. Now, Bohm proposes this provides a rough model that to some extent illustrates what an electron is and how it moves. We will consider this idea in more detail in the next section.

5. Enfoldment and unfoldment in quantum field theory

Bohm's radical suggestion is that a certain kind of enfoldment and unfoldment (which we have described in a simple and mechanical way above with the device) is actually taking place in the physical universe and is indeed the universal and fundamental feature of physical law. He points out that according to quantum field theory, all the elementary particles and atoms out of which chairs, tables, brains, bodies, and other physical objects consist are structures that are enfolded in principle throughout all space (1980, 209). Analogously to the visible droplets we see unfolding and enfolding in the glycerin tank model, what we call "particles" are momentary particle-like manifestations that constantly and very rapidly unfold and enfold from an implicate order that prevail in the movement of quantum fields. This is how Bohm himself describes the general picture:

> Because the implicate order is not static but basically dynamic in nature, in a constant process of change and development, I called its most general form the holomovement. All things found in the unfolded, explicate order emerge from the holomovement in which they are enfolded as potentialities and ultimately they fall back into it. They endure only for some time, and while they last, their existence is sustained in a constant process of unfoldment and re-enfoldment, which gives rise to

---

arises in the predictive processing framework is a question worth exploring in future research.

> their relatively stable and independent forms in the explicate order. (1990, 273.)

The enfolding-unfolding notion was inspired by a mathematical technicality, namely the Green's function approach to Schrödinger's equation. Feynman showed that in quantum mechanics one can use the Huygens construction to determine the wave function at a point y from the wave function at {x}, where {x) is the set of points on a surface at a previous time (Hiley and Peat 1987: 23). This can be expressed as:

$$\psi(y,t_2) = \int_{\text{surface}} M(x,y,t_1,t_2)\psi(x,t_1)\mathrm{d}x$$

Where $M(x,y,t_1,t_2)$ is a Green's function. The wave function at all points of the surface S contributes to the wave function at y. Thus, we can say the information on the surface S is enfolded into $\psi(y)$.

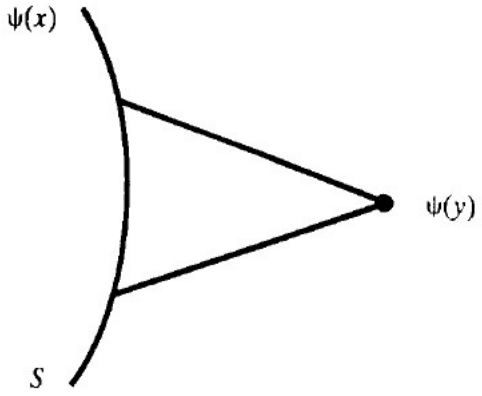

Figure 2. The information on the surface *S* is enfolded into ψ(*y*)

From this ψ(*y*) one can calculate the probability of finding the particle at the location *y*. Thus, we can say that the probability depends on the enfolded information of a set of earlier wave functions. In this way, via enfoldment, the probabilities that

prevail in the present moment in a given location depend on earlier wave functions.[10] The past constrains the probabilities of the present.

In turn, $\psi(y)$ itself gets 'unfolded' into a series of points on a later surface S'.

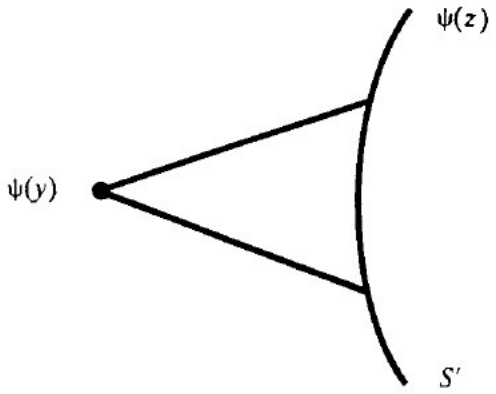

Figure 3. $\psi(y)$ gets 'unfolded' into a series of points on a later surface S'

Now, if we assume that the wave function is a complete description of the electron (as is often done in the usual interpretation of non-relativistic quantum theory) then the mode of being of an electron can be assumed to be a movement of unfoldment and enfoldment of the wave function.

Bohm and Hiley discuss the way the implicate order and quantum field theory are related in the last chapter of their 1993 book *The Undivided Universe*, noting that the Huygens' construction exemplifies the implicate order and is the basis

[10] In Bohm and Hiley's (1987) ontological interpretation of quantum theory, based on Bohm's 1952 "hidden variable" theory (which we are not discussing in this paper), the $\psi(y)$ determines the so-called quantum potential acting on the particle at y so that the particle reacts to the enfolded information of a set of earlier wave functions. So, in this model not just the probability but also the actual behaviour of the quantum object depends in a holistic way (via enfoldment) upon earlier wave functions. See Hiley and Peat 1987, 23.

of the Feynman graphs which are widely used. They point out that all matter is now analyzed in terms of quantum fields and note that the movements of all these fields are expressed in terms of propagators (used to calculate probability amplitudes for particle interactions using Feynman diagrams.) Thus, they argue that current physics implies that the implicate order is universal (Bohm and Hiley 1993, 355–356).

To make clearer the connection between the implicate order and Feynman diagrams, they use the glycerin tank model (see above, sections 3 and 4) to illustrate pair production and annihilation (see Figure 4). To model pair production, one needs to put in the droplets in such a way that when one turns the cylinder in a certain direction, it will appear that two particles emerge at a single region, and then start moving in different directions. To model annihilation, one needs to put in the droplets in such a way that when one turns the cylinder in a certain direction, two particles appear to come toward the same point from different directions, and then disappear.

Bohm and Hiley (1993, 359) describe how to do this. One first enfolds the droplets one by one, placing them at a constant distance from each other (marked by N in Figure 4). One then reverses the movement of the cylinder to enfold additional droplets (marked by S in Figure 4). Finally, one turns the cylinder in the original direction and then enfolds further droplets in a similar way. When the cylinder is then turned backwards, one will see a "particle" moving from the left toward point Q. At the point P there will suddenly appear a pair of particles moving in different directions and one of these will proceed to meet the original particle at the point Q. They then appear to annihilate each other, while the other member of the original pair will proceed onwards to the left (N).

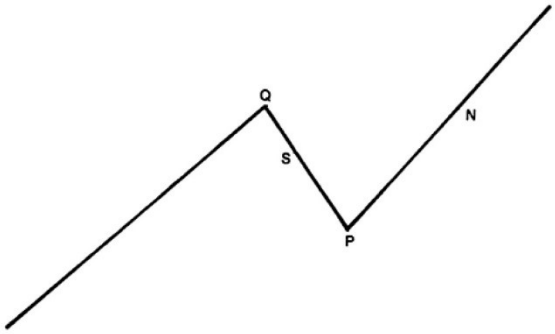

Figure 4. Enfoldment and pair creation.

Let us compare the above example to a Feynman diagram, which describes a somewhat similar situation. In the Feynman diagram shown below, an electron ($\mathbf{e^-}$) and a positron ($\mathbf{e^+}$) annihilate, producing a photon ($\boldsymbol{\gamma}$, represented by the sine wave) that becomes a quark-antiquark pair (quark **q**, antiquark $\bar{\mathbf{q}}$), after which the antiquark radiates a gluon (**g**, represented by the helix). Note that the positron ($\mathbf{e^+}$) and the antiquark ($\bar{\mathbf{q}}$) are moving backwards in time!

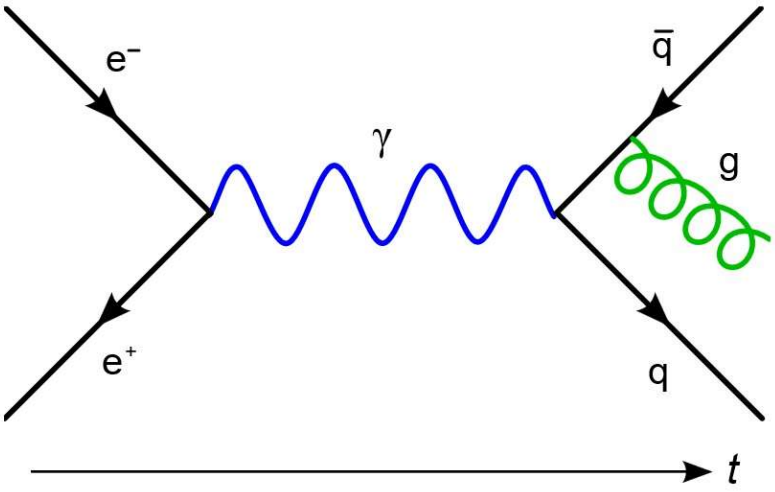

Figure 5. A Feynman diagram showing the radiation of a gluon when an electron and positron are annihilated. (Source: *Wikipedia*.)

Note, however, that unlike the representation given in the Feynman diagram, Bohm and Hiley do not say in their example that anything "moves" backward in time. Rather, they emphasize that what is significant is the order in the degree of implication.

They define an implication parameter, τ , of a droplet, which is proportional to the number of times the cylinder has been turned since that droplet was inserted:

> This implication parameter takes negative values when the cylinder is turned in the opposite direction. What happens in this example is that the implication parameter has a part that increases, another part in which it decreases and a third part in which it increases again. The entire pattern is present at each instant in the whole fluid with varying degrees of implication. All that happens with the passage of time is a change in the implication parameter which may be positive or negative. (Ibid., 360)

They then suggest that what is called the time coordinate in the Feynman approach may actually be the degree of implication:

> In this interpretation, Feynman diagrams would not refer to actual processes but rather to structures in the implicate order. The meaning of time would then be something different from τ but nevertheless related to it. (Ibid., 360)

So, the key point is that to explain particle creation and annihilation, it is not necessary to assume that particles move backwards in time. Physical processes involve implication and explication, enfoldment and unfoldment, but this does not require that anything moves backwards in time. What is characteristic of each "elementary particle" is its degree of implication, which, as time passes, determines the way it will appear in experimental situations, and how it appears to interact with other particles. Bohm and Hiley's model, in a sense, saves our usual notion of time, and draws attention to a new basic feature, unfoldment and enfoldment, which can be parametrised.

6. Concluding remarks

We saw above that there is a conservative element to Bohm and Hiley's approach. While Feynman is happy to propose radically that the positron ($\mathbf{e^+}$) and the antiquark ($\mathbf{\bar{q}}$) are mov-

ing backwards in time, this conclusion can be avoided in implicate order with its notion of enfoldment.

Note also the structuralist spirit of Bohm and Hiley'' discussion. They draw attention to the ""ntire pattern""that is present at each instant in the whole fluid with varying degrees of implication. And unlike what is commonly thought, they suggest that Feynman diagrams do not describe actual processes but rather structures in the implicate order. In another context, Bohm would characterize this structural aspect as the ""eing of becoming""(1986, 185; 197). While he is happy to assume that there is genuine becoming (i.e., ""ecoming of being"", he also assumes that there is an underlying timeless structure that is essential to becoming (this is the ""eing of becoming"". The above way of describing quantum field theoretical processes in structuralist terms is similar to Bohm'' description of our experience of movement when watching a motion picture (see section 4 above). While our experience is that of continuous movement, Bohm explains it structurally as grounded in ""o-present elements at different degrees of enfoldment""in the brain. Thus, whenever there is something that is ""parkling""(i.e., whenever there is movement in conscious experience or in the physical world), there is also something that is ""till""(e.g., the structures in the implicate order that make movement possible).[11]

Figure credits: Cindy Tavernise: Figure 1; Basil J. Hiley: Figures 2, 3 and 4; *Wikipedia* (public domain): Figure 5.

University of Helsinki and University of Skövde

References

Bohm, David (1952a & b). "A Suggested Interpretation of the Quantum Theory in Terms of Hidden Variables I & II". *Physical Review* 85, 66–179 and 180–193.

Bohm, David (1980). *Wholeness and the Implicate Order*. London: Routledge.

[11] I have borrowed the expression "Still or sparkling?" and part of the title of this paper from the title of section 3.7.1 in Ladyman & Ross' (2007) book *Everything Must Go: Metaphysics Naturalized*, where they discuss the metaphysics of time. Chapter 3 of that book is an excellent discussion of issues in philosophy of physics, even though the possibility of the individuality of quantum objects is dismissed a little too strongly. See Pylkkänen et al. (2016).

Bohm, David (1986). "Time, the implicate order, and pre-space". In *Physics and the Ultimate Significance of Time: Bohm, Prigogine and Process Philosophy*, edited by Griffin, 177–208.

Bohm, David (1987). "The Implicate Order and Prigogine's notion of irreversibility." *Foundations of Physics* 17, 667–677.

Bohm, David (1990). "A New Theory of the Relationship of Mind and Matter." *Philosophical Psychology* 3(2), 271–286.

Bohm, David and B. J. Hiley (1987). "An Ontological Interpretation of Quantum Theory." *Physics Reports* 144(6), 321–375.

Bohm, David and B. J. Hiley. (1993). *The Undivided Universe: An Ontological Interpretation of Quantum Theory*. Routledge: London.

Clark, Andy. (2013). "Expecting the World: Perception, Prediction, and the Origins of Human Knowledge." *The Journal of Philosophy* CX, 9, 469–496.

Dainton, Barry (2022). "Temporal Consciousness." *The Stanford Encyclopedia of Philosophy* (Summer 2022 Edition), edited by Edward N. Zalta. URL = https://plato.stanford.edu/archives/sum2022/entries/consciousness-temporal/

Freire, Olival Jr. (2019). *David Bohm, a life dedicated to understanding the quantum world.* Berlin: Springer.

Griffin, David Ray, ed. (1986). *Physics and the Ultimate Significance of Time: Bohm, Prigogine and Process Philosophy*. Albany: State University of New York Press.

Goldstein, Sheldon (2021). "Bohmian Mechanics." *The Stanford Encyclopedia of Philosophy* (Fall 2021 Edition), edited by Edward N. Zalta, URL = https://plato.stanford.edu/archives/fall2021/entries/qm-bohm/

Guay, Alexandre and Thomas Pradeu eds. (2016). *Individuals Across the Sciences.* Oxford: Oxford University Press. http://arxiv.org/abs/1405.4772

Hautamäki, Antti (2021). "Aika ja aikatietoisuus." ["Time and time consciousness"] In *Unfolding the Big Picture: Essays in Honour of Paavo Pylkkänen,* edited by P. Limnell ja T. Vadén.

Hiley, Basil J. and F. David Peat (1987). "General introduction: The development of David Bohm's ideas from the plasma to the implicate order". In *Quantum Implications: Essays in Honour of David Bohm*, edited by B.J. Hiley and F.D. Peat.

Hiley, Basil J. and F. David Peat, eds. (1987). *Quantum Implications: Essays in Honour of David Bohm.* London: Routledge.

Korpela, Sampsa (2022). *God, Time and the Concept of Potentiality in Quantum Physics.* Doctoral dissertation, University of Helsinki. Available at: https://helda.helsinki.fi/handle/10138/346531

Ladyman, James and Ross, D. (with D. Spurrett and J. Collier) (2007). *Every Thing Must Go: Metaphysics Naturalized.* Oxford: Oxford University Press.

Limnell, Petteri (2008). "*HIC & NUNC* – Nyt-hetken analyysia fysiikan, metafysiikan ja tajunnantutkimuksen näkökulmista." ["*HIC & NUNC* – Analysis of the moment now from the perspectives of physics, metaphysics and consciousness studies."] *Pro gradu* (Masters) thesis, University of Tampere.

Limnell, Petteri ja Tere Vadén, eds. (2021). *Unfolding the Big Picture: Essays in Honour of Paavo Pylkkänen*. Helsinki: Philosophical Studies from the University of Helsinki 50. English version: https://www.researchgate.net/publication/359867312_Time_and_Time_Consciousness_2

Marlow, A. R., ed. (1990). *Quantum Theory and Gravitation*. New York: Academic Press.

Pylkkänen, Paavo (2007). *Mind, Matter and the Implicate Order*. Berlin: Springer.

Pylkkänen, Paavo, Hiley, B. J. and Pättiniemi, I. (2016). "Bohm's approach and individuality." In *Individuals Across the Sciences,* edited by A. Guay, A. and T. Pradeu, 226–249. http://arxiv.org/abs/1405.4772

Wheeler, John Archibald (1980). "Pregeometry: Motivations and Prospects." In *Quantum Theory and Gravitation*, edited by A. R. Marlow, 1–11.

Wheeler, John Archibald (1990). "Information, Physics and Quantum: The Search for Links." In *Complexity, Entropy and the Physics of Information,* edited by W.H. Zurek.

Zurek, W. H. ed. (1990). *Complexity, Entropy and the Physics of Information,* edited by W.H. Zurek. Redwood City, California: Addison-Wesley.

Learning from the Past to the Future in Metaphysics

JANI HAKKARAINEN

1. Introduction: Engaging with the Past

This paper serves as a case study highlighting the significance of philosophical inquiry in actively engaging with its history, including its recent history. By "engaging," I mean the sincere consideration of the positions and arguments put forth by philosophers of the past, as well as learning from examining historical responses to these arguments. Such engagement does not entail rigidly adhering to canonical narratives of the (recent) history of philosophy or confining oneself to a specific philosophical tradition. Recognizing the intricate nature of the history of philosophy, which encompasses diverse perspectives and traditions, proves essential in uncovering a rich treasury of philosophical insights. The narratives we construct regarding the history of philosophy hold considerable importance for systematic philosophy.

In this paper, my focus is on a traditionally central field of philosophy: metaphysics. I will discuss the general metaphysical doctrine of Metaphysical Realism and especially its relevance to the discussion about the object of consideration and legitimacy of metaphysics.

The return of metaphysics in analytic philosophy since the 1970's coincides with the realist turn in analytic philosophy chiefly due to the work by Saul Kripke and Hilary Putnam in the first half of the 1970's and the rise of externalist epistemology from then on (e.g., Alvin Goldman and Robert Nozick). Since the 70's, there has been a vast amount of literature about realism(s). One of its forms discerned in the literature is *Metaphysical Realism* (Miller 2022). Typically, it is considered a jointly metaphysical and epistemological or se-

mantic doctrine (Miller 2022, Ch. 2). Considered as such, Metaphysical Realism is not only the claim that there is at least one mind-independent entity. Rather, it also involves the semantic or epistemic component that roughly, "we can meaningfully or truthfully say something about (the nature of) the mind-independent world" (Miller 2022, 23).

It seems to me, however, that Metaphysical Realism proper needs to be distinguished from epistemological and semantic realism since metaphysics is the study of being in so far as it is being (see below), whereas epistemology concerns epistemic states (e.g., understanding and knowledge) and philosophical semantics is about truth, meaning, and reference (among other things). Metaphysics, epistemology, and semantics are distinct fields of philosophy.

Therefore, in this paper, I follow a non-standard *working* definition of Metaphysical Realism as the ontological doctrine that *there are ontologically mind-independent entities standing in ontologically mind-independent relations*, such as numerical identity and distinctness, perhaps. One might think, for example, that there are numerically distinct physical objects that would be there even if there were no minds (e.g., the sun). Metaphysical Realism is widely held in one form or another among influential metaphysicians nowadays (e.g., Armstrong 1997, Lowe 2006, Fine 2009, and Schaffer 2009).

Accordingly, the *working* definition of the second-order view of *Metaphysical Realism about Metaphysics* (MRM), in turn, is that *metaphysics investigates metaphysically real entities and their metaphysically real relations* (whatever they are). As an illustration, E.J. Lowe defends the view that metaphysics investigates metaphysically real entities in their ontological categories, such as substances and kinds, that are not dependent on our categorizations (Lowe 2006, 195ff.). MRM presupposes Metaphysical Realism, since metaphysics cannot have a metaphysically real object of consideration if there are no metaphysically real entities. Given the popularity of Metaphysical Realism, it is only expected that MRM is also widely assumed or at least sympathized nowadays, by the metaphysicians mentioned above, for instance (see especially Lowe 2006 and Fine 2009).

These working definitions of Metaphysical Realism and MRM are intentionally tentative general formulations since

one of my main points is that it is the job of metaphysicians and metametaphysicians to make them more precise and to discuss competing more exact formulations of Metaphysical Realism and MRM with their respective merits, demerits, and relations to other relevant forms of realism, such as transcendental realism discussed critically by Kant.

Since antiquity, the epistemic value of metaphysical assumptions and assertions has been challenged in the *skeptical tradition*, especially if one assumes something along the lines of MRM. The skeptical tradition encompasses, for example, ancient Pyrrhonists, Francisco Sanches (1551–1623), and Hume. On a reasonable reading, Hume, for instance, advances a battery of skeptical arguments to the result that metaphysically real beings transcend the limits of the human understanding that are to be respected in philosophy (Hakkarainen 2012). Furthermore, Kant's critique of metaphysics as a *science* about transcendentally real beings is well-known, targeting especially Christian Wolff's (1679–1754) and Alexander Gottlieb Baumgarten's (1714–1762) ontologies and special metaphysics (cosmologies, rational psychologies, and natural theologies) (See Grier 2022).

The *typical* strategic move or just an (implicit) regulating assumption in the contemporary analytic metaphysics, more or less, is to bypass these critiques. However, it seems to me this move is not the best exemplar of intellectual integrity since the skeptics and Kant have many arguments that need to be engaged with.

As a start for the response to them from a metaphysician's point of view, I propose that metaphysical study is *initially* indifferent to the truth of MRM and Metaphysical Realism and does not presuppose them. Metaphysical Realism, as it is understood here, is a metaphysical doctrine the truth of which cannot be settled logically prior to metaphysical investigation. MRM presupposes Metaphysical Realism and therefore, one should not hold MRM uncritically. An *epistemological* consequence of this is that arguments against the possibility of cognition about metaphysically real entities (by e.g., Hume) are not arguments against the epistemic legitimacy of metaphysics without further arguments for MRM.

This paper is also a case study of the philosophical significance of having rich enough acquaintance with the history of

philosophy, the recent history included. I begin by discussing recent history in the next section and then proceed to what we can learn from it in the concluding third section.

2. Learning from the Near Past

To argue for my thesis about learning from the past in metaphysics, we better look at things from a different angle than usually in analytic metaphysics. Let us learn then from the phenomenological *formal ontological* tradition in metaphysics and ontology—that is, from the recent history of philosophy. Another insight can be learned from one of the most influential German philosophers of his time, the critical realist Nicolai Hartmann (1882–1950), who also engaged with phenomenology even though he was not a phenomenologist. Both may be seen as reactions to the Neo-Kantianism of the late 19th and early 20th century, as part of the German "new ontology" after the First World War; the 1920's and 1930's was not only the anti-metaphysics by, say, the logical positivists, such as Rudolf Carnap (1891–1970).[1]

The main representatives of the phenomenological formal ontological tradition are naturally Edmund Husserl (1859–1938), who coined the term "formal ontology" in his *Logical Investigations* (1900–1901) (Husserl 1970), and his pupils Edith Stein and Roman Ingarden (1893–1970). It has roots in the 19th century German Aristotle renaissance, too, especially in one of Husserl's professors, Franz Brentano (1838–1917) and Brentano's teacher Adolf Trendelenburg (1802–1872) (Albertazzi 2006, 43; Hartung, King, and Rapp 2019, 2–4). Together

[1] This is only one instance of the need for a rich enough (recent) history of philosophy. A new ontology evolved in Germany in the 1920's after the heyday of Neo-Kantianism: e.g., Hedwig Conrad-Martius (1888–1966), Hartmann, Max Scheler (1874–1928), Edith Stein (1891–1942), and Martin Heidegger (1889–1976) (Peterson 2019: xvii). In his *Ontology: Laying the Foundations* (*Zur Grundlegung der Ontologie*, 1935), Hartmann (2019, 3) also mentions Alexius von Meinong's (1853–1920) theory of objects (*Gegenstandstheorie*) in this context even though it was developed before World War I when Neo-Kantianism still dominated philosophy in Germany. However, Meinong was born and spent his professional career in Austria-Hungary, where Neo-Kantianism did not have the same position (Damböck 2020, 173).

with Kant and Hegel, this tradition created much interest in ontological categories (Albertazzi 2006, 53–54), which formal ontologists consider in a particular manner, as will be seen just below.

For the present purposes, Ingarden's project is the most relevant among the phenomenologists. He was a Polish philosopher trained in Poland and Germany. Ingarden's magnum opus is the three-volume *Controversy over the Existence of the World*, the two first volumes of which were written during the Second World War and published in Polish in 1947–1948, respectively (*Spór o istnienie świata*). Later Ingarden reproduced these in German and *Der Streit um die Existenz der Welt* was published in 1964–1965. He did not finish the third volume. Its incomplete version came out in Germany in 1974 (*Über die Kausale Struktur der Realen Welt*). Finally, the first two volumes were translated into English in Ingarden (2013) and (2016). In the *Controversy*, the fundamental question is how to put the distinction between idealism and realism precisely (Simons 2005, 39). This was motivated by Ingarden's view, right or wrong, that Husserl converted to transcendental idealism from transcendental realism just before the First World War (ibid.).

The *Controversy* is a highly intricate work and there is no need to go into all its complexities here. In this paper I will focus only on a few key points. In contrast to Meinong, for instance, Ingarden does not make any distinction between existence and being (Simons 1992, 378). Regarding the latter, he is a representative of the *mode of being* tradition in metaphysics, as well as Meinong and many others. Ingarden thinks that being is literally modified rather than unitary: there is a plurality of ways to be. (Ibid.) He does not think though that modes of being are primitive (Chrudzimski 2015). Rather, the possible modes of being, for example, the mode of being of intentional objects, need to be given an exact analysis by the different aspects of these modes of being, that is, *aspects of existence*, such as various ontological dependences (Chrudzimski 2015 and Millière 2016). Among the

modes of being, those of *being ideally* and *being really*[2] call for an analysis prior to trying to settle the issue of whether idealism or realism in any form is true (Chrudzimski 2015 and Millière 2016).

In Ingarden's case, it is crucial to appreciate that his conception of ontology is Husserlian. For him, following Husserl, ontology is the *a priori* discipline of what is essential to *possible* objects in general (Simons 2005, 40–1). Ontology differs from metaphysics, which investigates *what is*: for example, whether there are real beings (ibid.). Therefore, Ingarden's ontological analysis concerns primarily possible rather than actual modes of being.

Accordingly, Ingarden thinks that *being really* calls for an ontological analysis prior to trying to settle the metaphysical question about the existence of real beings. To be precise, the analysis of the mode of being of *being really* by aspects of existence belongs to a sub-field of ontology that Ingarden calls "existential ontology" (from "aspects of existence"), to which the first volume of the *Controversy* is devoted. Existential ontology differs from *formal ontology*, which Ingarden discusses in the second volume. Formal ontology investigates possible ontological categories (e.g., properties and processes), which differ from modes of being in Ingarden's view. He does not take ontological categories as primitive either. Possible ontological categories are analyzed first by modes of being and then ultimately in terms of aspects of existence. (Chrudzimski 2015 and Millière 2016)

As the present-day formal ontologist Peter Simons (2005, 41) has pointed out, Ingarden's three-partite distinction between ontological categories, modes of being, and aspects of existence can be reduced to the distinction between ontological categories and the ways in which entities exist. Following this reduction, I can say that the basic idea of *formal ontology* as a main branch of metaphysics is that it investigates *ontological categories* analyzing them by the ways in which entities exist—that is, by forms of being or *ontological forms* (Hakkarainen and Keinänen 2023, Chs. 1 and 3). In formal

[2] Here I write "being really" adverbially instead of "being real" in order to highlight the point that modes of being are ways of being literally modifying being.

ontology, existential dependence, for instance, is construed as an ontological form: existing dependent*ly*.

This does not entail, however, that ontological forms are modes of being, which would mean literal modifications of being. Without going into unnecessary details, I have argued that ontological forms are internal relations of a specific type called "formal ontological relations" in which entities stand (Hakkarainen and Keinänen 2023, Ch. 4). Entities are in one and the same sense, but they stand in distinct formal ontological relations. Here, the relevant point is that one can be a formal ontologist without being a mode of being theorist. (Ibid.) Formal ontology does not commit one to literal modes of being. Regarding ontological categories, we can take substances as an example of their analysis by ontological forms. An entity may be considered a member of the category of substances if the entity exists in the ontologically independent, numerically identical, persisting, and property-bearing way.

In accordance with this, I have proposed that it is illuminating to make a three-partite distinction between formal ontology, ontology, and general metaphysics (ibid., Ch. 5.3). To avoid a needless digression, I simply assume this distinction without arguing for it here. Concerning ontology, I follow more or less the mainstream Quinean line that ontology studies what there is (cf. metaphysics in Ingarden). This conception of ontology is reasonable, since "ontologia" is an early 17th century neologism coming from "ontos" and "logos" in Greek (see Lamanna 2014 and Smith 2022). "Ontos" is the possessive form of the Greek equivalent to "being" and "logos" means study or doctrine in this context. In the Quinean conception of ontology, "ontos" is understood in the *thing sense* of "being" meaning an entity (thing) or the totality of entities (every*thing*) (*ein Seiendes* or *das Seiende* in German and *olio* or *oleva* in Finnish).

However, "being" is ambiguous in English. Another sense of "being" is that which is shared by all entities: by definition, each and every one of them is there (in one way or another). Following English dictionaries, let me call this "the state sense of being": the state or condition of every entity is being

there (*das Sein* in German and *oleminen* in Finnish).[3] If "ontos" is taken in this sense of "being", then, in contrast to the Quinean conception, ontology is considered differently as the study of being in the state sense: *what it is to be*. This is rather the case in the German early modern conception of ontology as the science of being in general in Wolff and Baumgarten although their conceptions differ in detail (Lamanna 2014 and 2021). Hartmann agrees with it in his "fundamental question of ontology" "formally understood", as will be seen below (2019, 53). Above, I mentioned a third use of "ontology" in the Husserlian conception of it as the a priori science of the essence of possible objects as such. A fourth conception of ontology is the study of any object of thought and language in some historical conceptions (Jaroszyński 2018, 98 and Lamanna 2021).[4]

I supplement the Quinean conception of ontology as the study of what there is with the problem about the possible *ground of being* (Hakkarainen and Keinänen 2023, Ch. 5.3). This problem is not incompatible with the Quinean conception even though some metaphysicians think that the problem of ground is the ultimate question in metaphysics, rather than the Quinean existence question about what there is (e.g., Schaffer 2009). Therefore, I propose that the fundamental question of ontology in plain English is, *what is there and why*? Here "why" is roughly construed as asking the possible ground of being (in the thing sense) in virtue of which entities are there. Therefore, the fundamental ontological question is, to put it more precisely, what are entities and their ground(s)?

By contrast, formal ontology studies possible and actual ontological forms and ontological categories by analyzing categories in terms of ontological forms. General metaphysics, in turn, considers being in the state sense, its features, such as its unity/plurality and relation to existence and noth-

[3] It does not follow from this that being is a numerically distinct entity in the category of states, or, indeed in any ontological category. I use "state" here as a metaphysically neutral linguistic term. I also set aside the question how being is supposed to be expressed in language and thinking (whether by a quantifier or predicate, for instance).

[4] For a recent general history of the conceptions of metaphysics and ontology, see Jaroszyński 2018.

ingness, and the possible principles of being like the principle of non-contradiction. A crucial further point is that general metaphysics and formal ontology constitute a point of view from which ontology investigates entities and their ground. For example, if one asks the ontological question whether there are mind-independent universals, the framing of the problem presupposes something about being and the category of universals. (Ibid.)

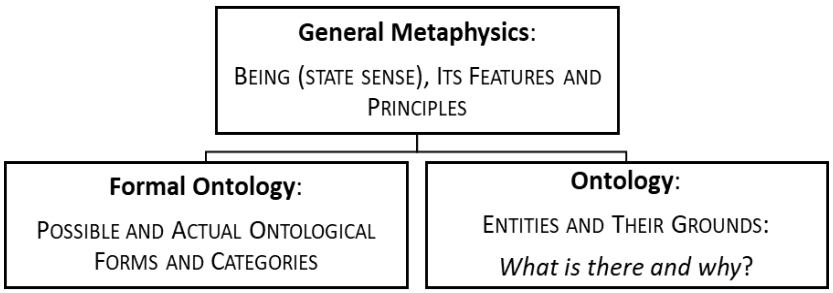

Figure 1. Main Branches of Metaphysics

Moving on to Hartmann, his insight about the beginnings of metaphysics is three-fold: (1) general metaphysics (in my terms) comes logically first and (2) begins with the question about being in the state sense, and (3) metaphysics is initially indifferent to the truth of Metaphysical Realism (again in my terms). We may learn his insight from his struggle with the epistemological Neo-Kantianism of his teachers Herman Cohen (1842–1918) and Paul Natorp (1854–1924) at his *alma mater* Marburg.

Especially illuminating is Hartmann's "problem of being" (*Seinsproblem* in German), the starting point of which, "the fundamental question of ontology", is Aristotle's (1958, 1003a 21) being as being (*to on hê on* in Greek: literally, 'that which is in so far as it is') (Hartmann 2019, 53). Hartmann explicitly distinguishes his problem of being from Heidegger's "question of being" (*Seinsfrage*) that asks the meaning of being (*Der Sinn von Sein*) (Hartmann 2019, 53–58). Hartmann thinks that "critical ontology" begins by asking what being *itself*, in the state sense (*das Sein selbts*), is (i.e., what it is to be), rather than the meaning of being. Critical ontology does not start with

the assumption or even the presupposition that being is dependent on or independent from, in one way or another, the subject cognizing it. In its beginnings, critical ontology considers being that is universally shared by all the beings and on which the cognizing subject does not impose any conditions. (ibid., 53–54) This is "the fundamental question of ontology" about being as being "formally understood" (ibid., 53).

Accordingly, Hartmann adopts Aristotle's "formula" of being as being for the reason that "because it considers what is [being in the thing sense] only insofar as it is, thus, only in its most universal aspect, it indirectly comes across 'being' [in the state sense: *Sein*] over and above 'what is' nonetheless." (ibid., 53) In short, the problem of being, that is, what it is to be, is logically prior to considering what there is.[5] I agree with Hartmann.

Hartmann thinks that critical ontology and the problem of being are initially indifferent to realism and idealism, transcendental realism and idealism in particular, although he eventually favors realism undogmatically (Peterson 2012, 295–296; Hartmann 2019, 51–54). Hartmann does not take any stance on the subject-dependence or subject-independence of being in the state sense at the beginning of ontology and metaphysics.

3. Prospects for the future

To conclude, I shall argue briefly that Ingarden and Hartmann indeed have a point from which metaphysics can learn. Following Aristotle, metaphysics is the study of being in so far as it is being. Therefore, Metaphysical Realism is, indeed, a metaphysical doctrine since it makes a claim that there are beings of a certain highly general sort: metaphysically real beings. If there are such entities, their being is of the metaphysically real sort. As this is still imprecise, *being metaphysically real* needs to be made exacter in order to be able to discuss whether there are such entities. We need to have a sufficiently exact grasp of the type of entities the existence of which we are investigating.

---

[5] This holds also of Aristotle's *Metaphysics*, in Vasilis Politis' view (2004, 3).

Here formal ontology proves to be helpful. Being metaphysically real is a plausible candidate for an ontological form since its tentative characterization is being independently from the mind. It is then a formal ontological problem informing the ontological doctrine of Metaphysical Realism how independence is to be specified here, assuming *hypothetically* for the sake of consideration that there are ontologically mind-independent entities. Is ontological independence existential or essential, for instance, identity independence (see Tahko and Lowe 2020)? Is it rigid (specific) or, as it seems to be, non-rigid or generic (ibid.)? In principle, an entity like the sun can depend for its *existence* or *identity* on a distinct entity, *modally* or *essentially* in non-modal terms.[6] It might be, for instance, that metaphysically real entities do not depend for their non-modal essence on any mind. Equally, an entity, such as the sun, may depend for its existence on a *specific* entity or a *type* of entities. The denial of the latter non-rigid dependence seems to be at play in the formulation of Metaphysical Realism. But is it so?

Generally: *if* there are metaphysically real entities, what is the correct analysis of their ontological form? Most likely, this kind of analysis produces different formulations of Metaphysical Realism, which can be then discussed critically in metaphysics. Whatever the case may be, answering these questions concerning what it is to be metaphysically real are presupposed by the very ontological problem setting about Metaphysical Realism: concerning the existence of metaphysically real entities in their metaphysically real relations. Formal ontology comes logically first in relation to ontology, which is the doctrine of *formal ontology first* by me and Markku Keinänen (Hakkarainen and Keinänen 2023, Ch. 5).[7]

---

[6] To most metaphysicians nowadays, the non-modal conception of essence is familiar from Kit Fine (1994). E.J. Lowe characterises it as follows: "In short, the essence of something, *X*, is *what X is*, or *what it is to be X*." (Lowe 2018, 16) This characterisation does not involve any modal term. By contrast, the *modal notion of essence* is roughly the de re necessary properties of an entity.

[7] Temporal order is a different issue. Conducting formal ontological study need not *temporally precede* ontological investigations. Rather, they may well be studied as a part of a single process of addressing metaphysical problems.

Furthermore, general metaphysical considerations of being in the state sense, its features, unity/plurality, and principles are *initially* indifferent to the truth of Metaphysical Realism. Metaphysical Realism comes to play in metaphysics only if we have good enough grounds to specify that it is *metaphysically real* being we are considering. If one is justified *not* to do that specification in a metaphysical study, it is possible to conduct the study and not to assume Metaphysical Realism.

This has consequences for ontology: Metaphysical Realism is relevant only if ontological problems presuppose or address it. The perennial topic of the existence of ontologically mind-independent universals, for instance, is an instance of such a case that presupposes Metaphysical Realism. One cannot even set the problem of their existence without having a sufficiently precise grasp of what it is to be metaphysically real. Equally, formal ontological problem settings presuppose something about general metaphysical matters, such as whether being is modified or not.

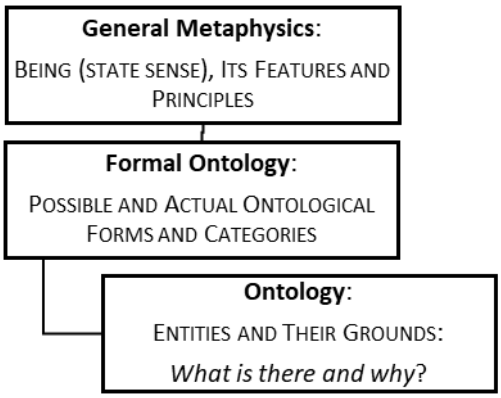

Figure 2. Main Branches of Metaphysics in Their Logical Hierarchy

Social ontology confirms that metaphysical study does not initially presuppose Metaphysical Realism. As a minimum, there are social ontological problems that concern possible entities that are ontologically *dependent* on the mind in one way or another, such as the existence and nature of some social kinds (e.g., gender and race) and institutions (e.g., money

and university).[8] Arguably, these entities would not be there if there were no minds; they depend for their existence non-rigidly or generically on minds, at least in modal terms. Still considering them is metaphysics since then one considers being in a manner that is closely informed by general metaphysics and formal ontology. Part of these problems is to make the type of ontological dependence exact.

Indeed, they presuppose Metaphysical Realism only if there can be nothing ontologically mind-dependent social in some relevant sense without there being something ontologically mind-independent in the same relevant sense. However, the truth of the latter proposition is an open metaphysical question not to be decided logically prior to investigation. We need to keep an open mind for the possibility along the idealist line that there are ontologically mind-dependent entities, but no ontologically mind-independent entities.

Regarding the *object of consideration* of metaphysics, the upshot is that without a justified specification to metaphysically real being, metaphysics is not a field studying such being in either the thing sense or the state sense; metaphysical antirealists can claim to study being. If it turns out that Metaphysical Realism is not justified, it does not follow that metaphysics cannot have any proper object or subject matter in some other reasonable sense. Metaphysical investigations may concern or even arrive at truths about entities although they are not *metaphysically* real, in accordance with my working hypothesis about Metaphysical Realism (e.g., some social entities).

In metaphysics, one should not assume the truth of Metaphysical Realism *uncritically*, still less implicitly. Metaphysical Realism is a metaphysical doctrine the justification of which is the task of the metaphysician. Neither can the age-old metaphysical controversy over realism and idealism be properly discussed without making its problem-setting exact. It calls for formal ontology discussing what the distinction between real and ideal entities is precisely. The relation between Metaphysical Realism and transcendental idealism by Kant or

[8] I consider only Metaphysical Realism here and set aside other forms of realism relevant to social ontology, since this is not a paper about that topic.

Husserl, for instance, is a complicated matter that calls for investigation. As was seen above, this is Ingarden's foremost motivation for his main work.

This upshot about Metaphysical Realism has consequences for MRM. Recall that MRM is the methodological second-order doctrine that can be given the following working characterization: the object of investigation of metaphysics consists in metaphysically real entities and their metaphysically real relations (e.g., numerical distinctness). MRM presupposes Metaphysical Realism, which needs to be justified by metaphysics. Therefore, it is not warranted to believe logically prior to metaphysical study that the object of consideration of metaphysics consists in metaphysically real entities; one must not put the cart before the horse. MRM is not to be assumed *dogmatically*, as Hartmann pointed out (in slightly different terms).

An epistemological consequence of all this is that arguments against the possibility of cognition of metaphysically real entities (e.g., Hume 2000 [1748], sec. 12) are not arguments against the epistemic legitimacy of metaphysics without further argument. They are so *only if* one makes a justified specification to Metaphysical Realism, that is, only against metaphysically realist metaphysical and metametaphysical views like MRM. This result indicates that our doctrine of *formal ontology first* provides a useful platform for discussing the epistemology of metaphysics, in addition to its other merits. That is, however, a prospect for the future study.

Acknowledgements

I would like to thank the following people very much for questions and comments: the two anonymous referees, the people present at my session in the Oulu colloquium, Alexander Carruth, Mirja Hartimo, Rögnvaldur Ingthorsson, Markku Keinänen, Anssi Korhonen, Vili Lähteenmäki, Sanna Mattila, Jaakko Reinikainen, Jenni Rytilä, and the students of the formal ontology seminar at Tampere University in the Spring term of 2022, especially Maarit Heikura, Matias Moisio, Thelma Nylund, Tupuna Wäljas, and Henrik Villanen.

Tampere University

References

Albertazzi, Liliana (2006). *Immanent Realism: An Introduction to Brentano*. Springer: Dordrecht.

Aristotle (1958). *Metaphysica*, edited by W. Jaeger. Clarendon Press: Oxford.

Armstrong, David (1997). *A World of States of Affairs.* Cambridge University Press: Cambridge.

Carruth, Alexander, Sophie Gibb, and John Heil, eds. (2014). *Ontology, Modality, and Mind*. Oxford University Press: Oxford.

Chalmers, David J., David Manley, and Ryan Wasserman, eds. (2009). *Metametaphysics: New Essays on the Foundations of Ontology*. Oxford University Press. Oxford.

Chrudzimski, Arkadiusz (2015). "Ingarden on Modes of Being." In *Objects and Pseudo-Objects: Ontological Deserts and Jungles from Brentano to Carnap*, edited by Seron, Richard and Leclercq, 199–222.

Chrudzimski, Arkadiusz, ed. (2015). *Existence, Culture, and Persons: The Ontology of Roman Ingarden*. Ontos: Frankfurt.

Damböck, Christian (2020). "(Dis-) Similarities: Remarks on "Austrian" and "German" Philosophy in the Nineteenth Century." In *Franz Brentano and Austrian Philosophy*, edited by Fisette and Stadler, 169–180.

Fine, Kit (1994). "Essence and modality." *Philosophical Perspectives* 8, 1–16.

Fine, Kit (2009). "The question of ontology." In *Metametaphysics: New Essays on the Foundations of Ontology*, edited by D. Chalmers, D. Manley and R. Wasserman, 157–177. Oxford University Press: Oxford.

Fisette, D. and G. F. F. Stadler (2020). *Franz Brentano and Austrian Philosophy*. Springer: Cham.

Grier, Michelle (2022). "Kant's Critique of Metaphysics." In *The Stanford Encyclopedia of Philosophy* (Fall 2022 Edition), edited by E. N. Zalta and U. Nodelman.
URL=<https://plato.stanford.edu/archives/fall2022/entries/kant-metaphysics/>.

Hakkarainen, Jani (2012). "Hume's Scepticism and Realism." *British Journal for the History of Philosophy* 20 (2), 283–309.

Hakkarainen, Jani and Keinänen, Markku (2023). *Formal Ontology*, Cambridge University Press: Cambridge.

Hartmann, Nicola (2019). *Ontology: Laying the Foundations.* Translated by K. Peterson. De Gruyter: Boston.

Hartung, Gerald, Colin Guthrie King, and Christof Rapp (2018). "Introduction: Contours of Aristotelian Studies in the 19th Century." In *Aris-*

totelian Studies in 19th Century Philosophy, edited by G. Hartung, C. G. King and C. Rapp, 1–10. De Gruyter: Boston.

Hume, David (1748/2000). *An Enquiry Concerning Human Understanding: A Critical Edition*. Edited by T. Beauchamp. Clarendon Press: Oxford.

Husserl, Edmund (1970). *Logical Investigations*. Translated by J. N. Findlay, Routledge: London.

Jaroszyński, Piotr (2018). *Metaphysics or ontology?* Translated by Hugh McDonald. Brill: Leiden.

Ingarden, Roman (2013). *Controversy Over the Existence of the World: Volume I*. Translated by Arthur Szylewicz. Peter Lang: New York.

Ingarden, Roman (2016). *Controversy Over the Existence of the World: Volume II*. Translated by Arthur Szylewicz. Peter Lang: New York.

Lamanna, Marco (2014). "Ontology Between Goclenius and Suárez." In *Suárez's Metaphysics in its Historical and Systematic Context*, edited by Novák, 135–152.

Lamanna, Marco (2021). "Francisco Suárez's Ontology (Science of Being)." In *Conimbricenses.org Encyclopedia*, edited by M. Santiago de Carvalho and S. Guidi.
URL=<http://www.conimbricenses.org/encyclopedia/suarez-francisco-ontology-science/>.

Lowe, E. J. (2006). *The Four-Category Ontology: A Metaphysical Foundation for Natural Science*. Clarendon Press: Oxford.

Lowe. E. J. (2018). "Metaphysics as the Science of Essence." In *Ontology, Modality, and Mind*, edited by Carruth, Gibb and Heil.

Miller, James T. M. (2022). *Metaphysical Realism and Anti-Realism*. Cambridge University Press: Cambridge.

Millière, Raphael (2016). "Ingarden's Combinatorial Analysis of The Realism-Idealism Controversy." In *Form(s) and Modes of Being. The Ontology of Roman Ingarden*, edited by Richard and Malherbe, 67–98.

Novák, Lukáš, ed. (2014). *Suárez's Metaphysics in its Historical and Systematic Context*. De Gruyter: Berlin.

Peterson, Keith R. (2019). "Translator's Introduction: Hartmann's Realist Ontology." In Hartmann (2019), xv–xxxix.

Politis, Vasilis (2004). *Routledge Philosophy Guidebook to Aristotle and the Metaphysics*. Routledge: London.

Richard, Sébastian and Olivier Malherbe (2016). *Form(s) and Modes of Being. The Ontology of Roman Ingarden*. Peter Lang: Bern and New York.

Schaffer, Jonathan (2009). "On What Grounds What." In *Metametaphysics: New Essays on the Foundations of Ontology*, edited by Chalmers, Manley, and Wasserman, 347–383.

Seron, Denis, Sebastien Richard and Bruno Leclercq, eds. (2015). *Objects and Pseudo-Objects: Ontological Deserts and Jungles from Brentano to Carnap*. De Gruyter: Boston.

Simons, Peter (1992). *Philosophy and Logic in Central Europe from Bolzano to Tarski*. Kluwer Academic Publishers: Dordrecht.

Simons, Peter (2005). "Ingarden and the Ontology of Dependence." In *Existence, Culture, and Persons: The Ontology of Roman Ingarden*, edited by A. Chrudzimski, 39–53.

Smith, Barry (2022). "The Birth of Ontology." *Journal of Knowledge Structures and Systems* 3 (1), 57–66.

Tahko, Tuomas E. and E. J. Lowe (2020). "Ontological Dependence." In *The Stanford Encyclopedia of Philosophy* (Fall 2020 Edition), edited by E. N. Zalta and U. Nodelman.
URL=<https://plato.stanford.edu/entries/dependence-ontological/>.

Can We Understand the Past? Wilhelm Dilthey on Historical Understanding

HENRIIKKA HANNULA

Wilhelm Dilthey (1833–1911) is today best known for his contributions to the philosophy of the human sciences and hermeneutics. In English, the term *human sciences* comprises both the social sciences and humanities, the German original is *Geisteswissenschaften* and the most literal Finnish translation would probably be *ihmistieteet* or *hengentieteet*.[1] Dilthey advocated for the independence of the human sciences from the natural sciences, arguing that the two were based on different methodological, epistemological, and ontological background theories. Dilthey's epistemology of the human sciences, and his epistemology of history, are built on this assumption of the special status of the human sciences vis-à-vis the natural sciences.

My aim in this paper is a systematic and concise step-by-step reconstruction of Dilthey's theory of historical under-

[1] The German term *Geisteswissenschaften* is notoriously difficult to translate as it carries specific philosophical and cultural-historical meanings not easily rendered into other languages. I will not dive deeper into this complex topic here; it suffices to say that the term designated a sphere of studies that had as their object the specifically human socio-cultural and historical existence and highlighted the special status of these studies vis-à-vis the natural sciences. In this paper, I have opted to use the term *human sciences*, which has been the established term in anglophone Dilthey-research after it was used in the translation of the *Selected Works I–VI*. For the reasoning, see Makkreel and Rodi (2019, vi–viii). In this paper, for clarity's sake, I have also opted to use the English translations of Dilthey's *Selected Works* (barring one reference, where translation was not available). I have, however, included the publication date of the German original, so that the reader can situate the discussions in their historical context.

standing. In other words, I try to figure out how Dilthey would have answered the question *of how we can know anything about the past, given that it no longer exists*. The first part of this paper will focus on the reconstruction of Dilthey's answers to the questions of what the human world is and how can we know anything about it. In the second part of the paper, I turn to *historical knowledge* and reconstruct Dilthey's answer to the question of how is it possible to know anything about the past. Dilthey's theory of historical understanding was, namely, built on his more general epistemology of the human sciences.

I first contextualize Dilthey's philosophical project in its intellectual-historical background and highlight its intellectual debt to German Historical School. Thus, I hope to make clear to the reader what kinds of debates Dilthey's philosophical project sought to contribute. I then identify the main epistemological and ontological arguments Dilthey used to ground human-scientific knowledge claims, concentrating on his ideas of the *volitional subject*, *lived experience*, *anthropological universalism*, and *understanding*. In the second part of the paper, I identify the conditions of possibility for historical understanding in Dilthey's theory. I identify *historicity* and *anthropological universalism* as such conditions. In the last part of the paper, I qualify Dilthey's theory by presenting and evaluating some popular objections that have been presented against Dilthey's theory of understanding.

Dilthey did not explicitly lay out his theory of historical understanding, even though it was at the heart of his philosophical project. This is why my main interest in this paper is exegetical. I, however, also want to highlight the virtues of Dilthey's theory and argue for the continuing relevance of his philosophy of the human sciences.

Dilthey and the Historical School – foundations of Dilthey's philosophical project

The development of Dilthey's philosophy and his conception of the human sciences must be understood against a specific historical background. The basic orientation of Dilthey's philosophical project was defined by the so-called *German Historical School* (Dilthey 1989 [1883], 48–49). The term Historical

School designated a varied group of thinkers from the late 18th to the early 20th century Germany, from different disciplines from historiography and linguistics to jurisprudence and economics.[2] Some important names were the jurist Friedrich Carl von Savigny (1779–1861), the historian Leopold von Ranke (1795–1886), and Gustav Schmoller (1838–1917) – the leader of the 'younger' German historical school of economics and best known for his involvement in the so-called method dispute with the Austrian Carl Menger. All these scholars were connected by the idea that *history* was the key source of knowledge for their respective fields and that historical research could provide its foundational insights.

The philosophical ideas underpinning the thought of the German Historical School have often been expressed under the concept of *historicism*. The concept has been defined in different, sometimes contradictory ways. Following Fredrick Beiser's definition, I understand historicism as the idea that "everything in the human world—culture, values, institutions, practices, rationality—is made by history, so that nothing has an eternal form, permanent essence, or constant identity which transcends historical change" (Beiser 2011, 2). Or, as Ernst Troeltsch defined it hundred years earlier, historicism was the "the fundamental historicization of all our thinking about man, his culture and his values" (Troeltsch 1961 [1922], 102). Nothing exists outside history, there are no meta-historical laws or principles, and the human world is constructed by its past.

The Schmoller-Menger debate is an especially illustrative example of the methodological commitments of the Historical School and those of its critics. It also well illustrates what was at stake in this controversy. Schmoller claimed that scientific knowledge in economics had to be based on empirical generalizations from historical material. Menger, in turn, maintained that economics needed to find universally valid laws and to posit ideal-typical explananda such as the *economizing individual* (see e.g., Dold and Rizzo 2023, 166–167; Louzek 2011, 440; Haller 2004, 17–18). Menger thereby wanted to model economics more after the natural sciences and criti-

[2] For a good introduction and overview of this tradition, see e.g., Beiser 2011, 1–23.

cized the non-exact nature of historical methodology vis-à-vis abstract logical principles. Schmoller eschewed the hypothetical abstractions involved in positing universal laws.

A similar tension between historical and abstract approaches resurfaced many times throughout the 19th century in different debates involving the members of the Historical School. Dilthey himself accused August Comte, John Stuart Mill, and Henry Thomas Buckle of illegitimately borrowing principles and methods from the natural sciences and thus over-simplifying and distorting the complex human reality into abstract principles and ideas. They attempted to "solve the riddle of the historical world by borrowing methods from the natural sciences", while the Historical School "could only protest against their impoverished, superficial, but analytically refined results by appealing to a more vital and profound intuition" (1989 [1883], 48–49, see also: 154). But a "profound intuition" alone cannot ground human sciences philosophically. In the preface to the programmatic 1883 *Introduction to the Human Sciences*, Dilthey motivates his philosophical project as follows:

> Even today the Historical School has not yet succeeded in breaking through the inner limits which have necessarily inhibited its theoretical development and its influence on life. Its study and evaluation of historical phenomena remain unconnected with the analysis of facts of consciousness; consequently, it has no grounding in the only knowledge which is ultimately secure; it has, in short, no philosophical foundation. (Dilthey 1989 [1883], 48.)

Dilthey himself set out to find such a secure foundation. There are two important takeaways from this quote. First, Dilthey's project was from the very beginning tied to this specific conception of the human sciences which came from the Historical School. Opposing views would, for example, have been the positivism of August Comte or the philosophy of history of G.W.F. Hegel, who both purported to be able to uncover meta- or ahistorical laws governing history and the social world (ibid. 153–156). For Dilthey, knowledge in the human sciences was in some important sense based on historical, immanent, and thus empirical description and analysis – as opposed to idealist or positivistic theoretical constructions.

Secondly, Dilthey's project sought to provide a philosophical foundation for this specific conception of the human sciences. But what does the somewhat elusive idea of a philosophical foundation mean? A popular line of interpretation has been to read Dilthey as *some sort* of Kantian: that the search for a philosophical foundation meant expanding Kant's primarily nature-oriented *Critique of Pure Reason* into a *Critique of Historical Reason* (see e.g., Holborn 1950, 99–100; Ermath 1978; Makkreel 1975, 8; 2021; Rockmore 2003, 577). Dilthey did occasionally refer to his philosophical project in these terms. His theory of history can indeed be reconstructed as looking for the conditions of possibility for historical knowledge and in this sense can be framed as a critique in the Kantian sense. However, Dilthey explicitly rejected the fundamental Kantian notion of a formal a priori. Dilthey's naturalized or historicized apriori could only be arrived at via an empirical and descriptive approach. Dilthey thereby positioned his theory in many ways explicitly against Kant and the neo-Kantianism of his own time. Therefore, I agree with for example Theodore Plantinga's reading that the term 'Kantian' can only really be applied to Dilthey if it is emptied of most of its meaning (1980, 18).

Dilthey's philosophical project was in many ways explicitly *anti*-Kantian. Instead of transcendental logic, Dilthey sought to provide a *psychological* and *anthropological genealogy* of human cognition. Psychology and anthropology would give philosophical legitimacy to the human sciences, not apriori logic. This goal clashed quite profoundly with orthodox Kantian critiques. Dilthey's philosophical foundation of the human sciences also included rethinking the very question of *what a knowing subject is*. He also needed to answer the question of *what makes understanding between people possible.* Furthermore, to bring the human sciences on a philosophically solid foundation, we also need to know what their object is. *What the human world is*, what constitutes sociocultural structures, and how do they relate to individual agency?

Rethinking the thinking subject

In the 1883 *Introduction to the Human Sciences*, Dilthey famously states that "no real blood flows in the veins of the knowing

subject constructed by Locke, Hume, and Kant, but rather the diluted extract of reason as a mere activity of thought" (Dilthey 1989 [1883], 9). In opposition to this abstract subject, Dilthey's *thinking human* being relates to the world not only through intellect or "mere activity of thought", but also through emotions and volitions.

Our cognition, our way of knowing the world, is thoroughly shaped by our sense of agency – what we want to achieve, what we cannot achieve, and how all of this makes us feel. For Dilthey, the human world is not a noumenal thing-in-itself that our cognition gives shape to. Rather, Dilthey claims that the human psyche and the surrounding world are in constant interaction with each other and thus co-constitute each other. Human cognition arises from a dynamic interaction with the world. When people seek to fulfill their goals and create things, they do this in the framework of the existing sociocultural world but at the same time construct the sociocultural world further.

Dilthey argued that thinking is shaped by "agency and suffering, action and reaction" (Dilthey 2002a [ca. 1900/1905], 219). Humans are not merely *subjects* who operate with *objects* or *ideas*, but our whole relationship with the world is colored by what we want and what we feel. When I am hungry, I do not just cognize the sandwich I see on the table: I also want to eat it and I am attuned to it accordingly. If I want to change something about the world or achieve my goals, I need to use force to enact change. I feel the pressure and resistance of the world as it prevents me from achieving my goals and I struggle against it (Dilthey 2010a [1890], 8–18). The individual's relationship to the world is defined by a constant struggle against facticity, the desire to move, change, and have things. Thus, the nexus of our beliefs about reality is *volitional*, first and foremost. The failure of classical philosophical epistemology as written by Kant, Locke, Hume etc. had been – Dilthey argued – that it treated representations and ideas as the be-all and end-all of human knowledge.

In some ways, Dilthey's theory of human sciences also anticipated later phenomenological philosophy. Dilthey took as his starting point the study of phenomena as they appear in consciousness – as meaningful, tangled with intentions and hopes, and thus also emotionally imbued. Husserl's later ide-

as about the significance of intentionality for phenomenology and Heidegger's theories about moods were thus in many ways prefigured in Dilthey's thought, and both philosophers were also directly influenced by Dilthey.[3] This fully-fledged lived experience – volitional, emotional, and rational – Dilthey argued, must be taken as the starting point of epistemological inquiry. Consequently, also the philosophical foundation of the human sciences must be built on this insight.

Anthropological universalism as the precondition of understanding

Every human being relates to the world through the same triad of *feeling, willing*, and *thinking*. Qua being humans, we share this same volitional structure of experience. However, given that throughout our individual development we go through different kinds of feelings and volitions, life molds us into different, unique individuals (Dilthey 2010b [1895–96], 211–212). This is the process of individuation. Yet, since the core potential in every person is the same, we end up sharing enough similarities that no one's inner life is completely and incommensurably alien to us. This makes mutual understanding between different individuals possible. Dilthey grounds the possibility of understanding between people on an idea of "immanent universalism" (see also: Kinzel 2019, 27–30) or, as I shall here call it, anthropological universalism. He describes the idea as follows:

> [A]ll humans arrive at the same numerical system, the same spatial, grammatical, and logical relations. Because they live in the midst of relations between this external world and a common structural psychic nexus, they share the same forms of preference and choice, the same relationships between ends and means, certain uniform relations of values, and certain uniform features characteristic of their ideal of life wherever it appears. (Dilthey 2010c [1894], 196.)

[3] For example, Makkreel (1982), Nelson (2014) and Scharff (2019) have written about different aspects of Dilthey's influence on Husserl and Heidegger.

Dilthey's universalism is thereby not based on essentialist assumptions about what it is to be a human being. If we would put a baby to grow up in a radically different kind of environment – to the Moon, for instance – without human contact, it might grow up to have a radically different way of cognizing the world and understanding would not be possible anymore. The possibility of mutual understanding hinges on the fact that we are certain kinds of beings in a certain kind of world.

Dilthey's unity in diversity-ethos and romantic holism might appear old-fashioned and even somewhat problematic today. In many ways, his universalism indeed owed to the theologically imbued ideas of people like Schleiermacher and Herder, and German romanticism (see e.g., Gadamer 2013 [1960], 226; Shalin 1986, 116–117). The legitimacy of this kind of universalism can of course be questioned: can we genuinely assume that people's experiences across different cultures or circumstances are truly similar enough to secure the possibility of mutual understanding? For example, one can plausibly argue that understanding someone's extreme suffering – say, of holocaust survivors, victims of war or colonial subjugation, or people experiencing racial discrimination – is not possible for someone who has not gone through those experiences (for background, see e.g., White & Wang 1995; Kidron 2012). Furthermore, it is also plausible that those experiences might be impossible to explain or communicate to others (see e.g., Kusch 2019). If properly understanding the experiences of my neighbor is sometimes impossible, how could I claim to understand someone from a radically different time or culture?

However, Dilthey's universalism might be defended if we take his conception of universality to be thin and elementary. Essentially, Dilthey argued for the existence of anthropological or cultural universals. The existence of such universals – elements, patterns, traits, or institution that is common to all known human cultures worldwide – is still debated in contemporary anthropology (see e.g., Pinxten 1975, 7–8; Lloyd 2007, 1–8) and therefore cannot be rejected outright. In this (more charitable) reading, Dilthey simply postulated that humans share just enough psychological similarities that make understanding other people *principally* possible. It

could even allow that certain lived experiences cannot be communicated or understood, while still maintaining that the *condition of possibility* for mutual understanding between people is that we are all in some significant ways similar.

Understanding, explanation, and *lived experience*

Dilthey's name is often associated with the so-called understanding vs. explanation distinction, whereby explaining is something that the natural sciences do, and understanding is the epistemic strategy of the human sciences. Based on this distinction, Dilthey maintains that the human sciences are inherently distinct from the natural sciences: the former concentrates on the understanding of meaningful human actions, while the latter depends on the causal explanation of physical events. Even though this idea did not originate from Dilthey (see e.g., Maclean 1982, 348–349), he provided perhaps the first properly systematic formulation of it. Dilthey's clearest and the most comprehensive formulation of the explanation-understanding-distinction is: "[w]e explain through purely intellectual processes, but we understand through the cooperation of all the powers of the mind activated by apprehension" (Dilthey 2010c [1894], 147).

Dilthey's conception of human-scientific understanding is directly based on his reconceptualization of the thinking subject and his anthropological universalism. Understanding in the human sciences is analogous to understanding of other people. As argued above, our subjectivity is not constituted only by rational deliberation, but also (and perhaps more profoundly) by volitional and emotional states. All these powers of the mind are activated when we try to understand other people and, therefore, also in human-scientific understanding. The human world is not a rigid object to be sensed or cognized, but a playing field in which human beings try to fulfill their projects and desires. Understanding this field requires engaging all human cognitive capacities. A scholar of the human sciences inhabits the realm that is also the object of her study. They are subject to the same psychological and anthropological realities that govern the rest of the human world. This is why the human sciences, unlike the natural sciences, are directly connected to life:

> Life, life-experience, and the human sciences are intimately related and constantly interact. The basis of the human sciences is not conceptualization but the reflexive awareness of a psychic state in its wholeness and its rediscovery in re-experiencing. (…) Here we notice a decisive difference between the natural sciences and the human sciences. In the former, scientific thinking stands apart from our ordinary contact with the external worlds (…) But in the human sciences, a connection between life and science is retained, so that the thought-provoking work of everyday life remains the foundation of scientific creativity. (Dilthey 2002b [1911], 157–158.)

In *understanding*, we start from the immediate givenness of our own *lived experience* (i.e., the experience of the word). However, this immediate givenness is conceptualized and articulated into a whole we can analyze and describe. Because of the dialogical relationship between the individual and other people, the sociocultural world is always already implicated in the lived experience.

Dilthey's theory of understanding sought to establish and ground a theory for *epistemic access* (Kinzel 2018, 356–360) to the human world. Cultural facts only exist if someone can experience them. There cannot be a cultural or a social fact that does not feature in at least someone's experience. There cannot be a norm, custom, or institution that nobody in the world knows about. This means that all human scientific facts are accessible through *lived experiences* of one or several individuals. The facts of the natural world, in contrast, exist irrespective of us. A species of bird that no human has ever seen still exists. Lived experience, as Dilthey understands it, is a window through which a researcher can peek into the vast and complex reality of the human world. Through *understanding*, we can access lived experiences that are not our own.

The other side of the coin – explanation – would explain the natural world from the outside by referring to causal explanations or explanations identifying covering laws. The epistemological difference implies an ontological difference as well – if understanding the sociocultural world is qualitatively different from explaining the natural world, there must be something about the ontological makeup of the sociocultural world that causes this difference. In more modern

terms, Dilthey's understanding vs. explanation distinction was based on a social-ontological theory. He argued that *social kinds* are different from *natural kinds*. Social kinds arise in the interaction between people and create the sociocultural world. Because we – as humans – co-constitute the latter, we have a different kind of epistemic access to that world than to the world of nature. Human life is structured and carried out through meaningful action and symbolic expressions. Consequently, the human world is constituted by manifestations of life, such as artworks, texts, words, utterances etc.

To recapitulate the argument this far: Dilthey's epistemology was supposed to provide a model of a historically and culturally situated mind. It did this by grounding our knowledge claims on lived experience. Furthermore, it maintained that we can have access to other people's inner states, or, in other words, that we can understand other people and their expressions of life. This understanding is qualitatively different from the causal explanations of the natural sciences. Dilthey's explanation-understanding distinction was thus based on a substantive epistemological and ontological background theory.

Dilthey on historical understanding

Until now, I have concentrated on Dilthey's theory of understanding in general. However, Dilthey's theory of epistemic access to the sociocultural world does not yet establish the possibility of historical understanding. We cannot directly experience the past, as it is no longer there. Hence, we cannot have a lived experience of the past. Dilthey goes around this hurdle by explaining that even though the past is no longer there, it is still implicated in our experience of the world:

> Everything here derives from acts of human spirit and bears the hallmark of historicity. As a product of history, everything gets interwoven with the world of sense. The distribution of trees in a park, the arrangement of houses in a street, the handy tool of the artisan, and the sentence propounded in the courtroom are everyday examples of how we are constantly surrounded by what has become historical. What the human spirit is today projecting into some manifestation will tomorrow, when it stands before us, be history. Through the passage of time we become

> surrounded by Roman ruins, cathedrals, and the country palaces of autocrats. History is not something separated from life or remote from the present. (Dilthey 2002b [1911], 169.)

As mentioned earlier, Dilthey subscribes to the idea of historicism, the philosophical idea that one takes things to be historically constituted in some important sense. It maintains that to explain the existence of social and cultural phenomena, one must study the process by which they came about. In other words, their history. Apart from the concrete ways in which the past is implicated in our material surroundings, Dilthey also describes the past being present as a kind of memory:

> because what we experience as present always contains a memory of what has just been present. In other cases, the past has a direct effect on, and meaning for, the present, and this gives to memories a peculiar character of being present through which they become included in the present. (ibid. 216.)

This past being present means in a very broad sense that the world in which we live is a product of history. The past is implicated in the present in the same way as someone's childhood is present in one's adult life. It is constitutive of what the world is today.

Dilthey understands historicity as a kind of living tradition that is carried forward in different social institutions and social configurations. The ontological historicity does not need to mean a spooky ethereal historical spirit hovering above the present. Rather, historical categories manifest themselves in the present through institutions, structures, ideas, and social configurations. One such structure would, for example, be *generation*. In fact, Dilthey was the first German scholar to use this key sociological concept systematically (Mannheim 1952 [1927/28], 282). In his 1871 biography of the philosopher Friedrich Schleiermacher, Dilthey demonstrates how the previous two generations before Schleiermacher created the intellectual world which made Schleiermacher's thought possible in the first place. He also accounts for how the following generation developed Schleiermacher's theories further and partly reacted against them. The past is anchored in concrete, empirical categories; institutions and the organiza-

tion of society are carriers of historicity. Historicity is expressed in real social kinds and is thus woven into the fabric of the social world.

Dilthey's theory of historical understanding, therefore, includes the general requirements of *lived experience* and epistemic access to the human world, or – in other words – *understanding*. It also includes *historicity* as a condition of possibility for historical understanding. *We can understand the sociocultural world because we are part of it, and we can understand historical categories because the world implicates them.*

However, to say anything meaningful about the past or, especially, to say anything legitimate about it as a historian, the historicity of the human sociocultural world does not bring us very far. It just establishes the fact that *some* historical categories are not foreign to us, as they are present in our own lives. But what about phenomena or ideas which are not implicated in our historical consciousness or manifested in our social and cultural world? What about historical and cultural distance? Many medieval institutions, like universities, still exist and carry the ideas and ideals directly to our lives, but this is not the case for, say, many ancient Greek cultural institutions like the Panathenaic Games or the boule. Their meaning and significance are lost to us, at least when it comes to our own lived reality.

To answer this question, Dilthey would again invoke his idea of *anthropological universals*. Historical understanding is possible because there are still enough commonalities between people of the past and the present. All people share the same basic structure of experience or basic attunement to the world – all people desire, love and hate. They create laws and art and create organizational structures and cultural systems to make these possible. This commonality can again be understood very minimally: almost all cultures everywhere have had some kind of conception of marriage, for instance. Conceptions of marriage have varied greatly across cultures and epochs, but people still seem to tend towards creating an institution to secure familial continuity. All social groups create laws and societal structures. Religion is ubiquitous, even though the details of religious formations vary. This very minimal and basic level of commonality secures the *condition*

of the possibility of understanding culturally or temporally remote manifestations of life.

Dilthey never claimed that such understanding would be easy. He admitted that understanding other people and their expressions of life gets more difficult when the distance increases: "the certainty of our knowledge rests in us ourselves, and as we move away from ourselves, our knowledge becomes gradually less certain" (Dilthey 1977 [1860], 208). Understanding comes in degrees from elementary to complex. Consequently, there might well be historical phenomena that cannot be understood, because not enough source material exists to make understanding possible. Understanding must be based on empirical material. Still, even the most complicated effort of historical knowledge is nevertheless related to and derivative from our knowledge of other persons and our own awareness of life:

> Understanding ranges from the babblings of children to Hamlet or the Critique of Pure Reason. From stones and marble, musical notes, gestures, words and letters, from actions, economic decrees and constitutions, the same human spirit addresses us and demands interpretation. Indeed, the process of understanding, insofar as it is determined by common conditions and epistemological instruments, must everywhere present the same characteristics. (Dilthey 1996 [1900], 232.)

The abstractions and analyzes of a historian are a critical extension of the processes of understanding which are implicit in our ordinary experience. The understanding of historical agents, individuals who lived in different times is analogous to understanding the living people around us. The difference between understanding past people and living people is only due to the difference in the manifestations of their life that are available for interpretation. Interpreting historical texts is maybe more challenging but not a fundamentally different task than reading a contemporary novel or understanding the complaints of my neighbor. As Dilthey puts it: "through stone and marble, musical notes, gestures, words, and texts, actions, economic regulations and constitutions, the same human spirit addresses us and demands interpretation" (ibid. 232–233).

Dilthey criticized and Dilthey defended

A popular later reading of Dilthey has (over-)emphasized the role of *empathy* in his theory of understanding. Already Dilthey's contemporary neo-Kantian philosopher, Heinrich Rickert, claimed in quite an uncharitable manner that "certainly Dilthey's gift for 'reliving' and 'empathizing' history was extraordinary and perhaps unique in its time. However, this estimable man was not gifted in the same measure with the capacity for rigorous conceptual reasoning" (1962 [1926], 22). The juxtaposition of soft and unsystematic hermeneutic understanding with rigorous and careful conceptual reasoning has been a recurring theme in Dilthey-reception since the beginning. It has also sent echoes more generally in the reception of hermeneutic methodologies in the human sciences, and the contrasting of 'soft' qualitative methods against the more 'exact' quantitative methods.

In this view, sympathy and understanding are understood as psychological mechanisms that allow re-enacting or reliving someone's psychic states – walking a mile in someone else's shoes, so to say. Historical understanding would thereby be more akin to imagining historical agents' mental states. To a certain extent, one must admit, Dilthey indeed champions such a psychological theory. However, his theory of understanding cannot be reduced to it. Dilthey never conceives of understanding *solely* as an act of empathetically 'transporting' oneself into the head of another person. As argued above, Dilthey's psychological universalism is merely the first *condition of possibility* for human-scientific and historical understanding. Interpretation is an active and rigorous process, based on empirical evidence. Human-scientific knowledge production also involves gathering empirical material, even by using quantitative and statistical methods, and critically evaluating it. Dilthey is also a spirited defender of modern source criticism. Historical understanding, for him, is therefore not primarily an act of imaginative empathy but a rule-bound interpretative process. It is thus disingenuous to reject Dilthey's conception of historical understanding by reducing it merely to empathy.

Another influential line of criticism of Dilthey's conception of understanding comes from later philosophical hermeneu-

tics, especially from Martin Heidegger and Hans Georg Gadamer. This difference between Dilthey vs. Heidegger and Gadamer has been conceptualized as one between *methodological* vs. *philosophical-ontological hermeneutics* (see e.g., Gjesdal 2017, 337; 2019, 358, 369; Fehér 2015, 162). Both Heidegger and Gadamer argue that Dilthey's conception of understanding tries to uncover the intended or original meaning of a phenomenon and is therefore misguided. Instead, the meaning or significance of a phenomenon should be understood as co-constituted by the person doing the understanding. For Heidegger, and later for Gadamer, one of the main problems with Dilthey's "methodological" hermeneutics was that it divorces the object of study – the object of understanding – from the person trying to understand it.

In *Being and Time*, the crux of Heidegger's criticism is existential: Dilthey's thing-like attitude towards history produces inauthentic historiography which does not consider the primordial existential significance of the historicity for *Dasein* (1977 [1927], 518–532). In a similar vein, Gadamer accuses romantic hermeneutics in general, and Dilthey in particular, of naïve historical objectivism (Gadamer 2013 [1960], 226). Dilthey's hermeneutics tried erroneously to reach the object *as it really was*. Romantic or methodological hermeneutics, Gadamer declares, falsely treats understanding as a one-way subject-object relation, when in fact understanding means the fusion of horizons, in which the old and the new horizon are combined into something of living value (ibid. 317). This also entails a different conception of historicity: Dilthey's idea of historicity was tied to his quest for the epistemic justification in the human sciences, while Heidegger and Gadamer saw it as an existential category that mediates meaning and tradition (Gjesdal 2014, 305–307). To simplify, if for Dilthey the past was something we want to understand scientifically, for Heidegger and Gadamer historicity and tradition disclosed meanings for human existence. In this reading, Dilthey's theory was epistemological while Heidegger's and Gadamer's were primarily existential and ontological.

The main motivation of Dilthey's philosophical project indeed is to prove the possibility of scientific understanding in the human sciences. The epistemological questions, of how we can know anything about the sociocultural world or the

past, are at the forefront of his philosophy. For some, this might exactly be the advantage of Dilthey's philosophy vis-à-vis Heidegger and Gadamer. However, as I have tried to show in this paper, Dilthey's philosophy of the human sciences is not entirely reducible to epistemology. His theory of historicity is also ontological, and his rethinking of the thinking subject likewise precludes accusations of naïve objectivism. The difference between Dilthey and Heidegger and Gadamer is less systematic than the sharp distinction between methodological vs. ontological-philosophical hermeneutics suggests. Rather, the difference comes down to their different philosophical interests and motivations, what kind of ideals guided their philosophical inquiry and what problems they were trying to answer.

Dilthey tried to find a philosophically solid foundation for the human sciences and to explain how human-scientific and historical understanding is possible. However, again, Dilthey's theory only really establishes the conditions of possibility for this understanding. The part and parcel of human-scientific or historical understanding comes after that, it is the skilful interpretive engagement with empirical material. As such, it is not obvious why Dilthey's theory would have to clash with the ontological hermeneutics of Heidegger and Gadamer. Apart from epistemology, Dilthey's philosophy of the human sciences also includes a sophisticated social and historical ontology and an interesting theory of human agency.

Acknowledgements

The earliest version of this paper was specifically written for the *Annual Colloquium of the Philosophical Society of Finland 2022*, organized in the University of Oulu 9–10 June 2022. I want to thank the participants of the colloquium for helpful comments and questions. I am likewise grateful for the participants of Martin Kusch's *Seminar for Doctoral Candidates* at the University of Vienna in winter semester 2022. Having had the opportunity to present this paper at both these events has helped me to develop and clarify its arguments greatly. Likewise, I thank two anonymous referees for their kind and constructive comments. I would also like to extend my sin-

cere thanks to the Eino Jutikkala Fund and the Emil Aaltonen Fund for making my research possible in the first place.

University of Vienna

References

Beiser, Frederick C. (2011). *The German Historicist Tradition*. Oxford: Oxford University Press.

Dilthey, Wilhelm (1977). "Anhang: Frühe Aphorismen aus Berliner Zeit (vor 1860)." In *Wilhelm Dilthey Gesammelte Schriften vol 18*, edited by Helmut Johach and Frithjof Rodi. Göttingen: Vandenhoeck & Ruprecht, 203–213. (Original work published in 1860)

Dilthey, Wilhelm (1989). "Introduction to the Human Sciences." In *Wilhelm Dilthey Selected Works I. Introduction to the Human Sciences*, edited by Rudolf Makkreel and Frithjof Rodi. Princeton, NJ: Princeton University Press, 47–219. (Original work published in 1883).

Dilthey, Wilhelm (1996). "The Rise of Hermeneutics." In *Wilhelm Dilthey Selected Works, Volume IV: Hermeneutics and the Study of History*, edited by Rudolf Makkreel and Frithjof Rodi. Princeton: Princeton University Press, 235–258. (Original work published in 1900).

Dilthey, Wilhelm (2002a). "The Formation of the Historical World in the Human Sciences." In *Wilhelm Dilthey Selected Works III. The Formation of the Historical World in the Human Sciences*, edited by Rudolf Makkreel and Frithjof Rodi. Princeton, NJ: Princeton University Press, 101–174. (Original work published in 1911).

Dilthey, Wilhelm (2002b). "Plan for the Continuation of the Formation of the Historical World in the Human Sciences." In *Wilhelm Dilthey Selected Works III. The Formation of the Historical World in the Human Sciences*, edited by Rudolf Makkreel and Frithjof Rodi. Princeton, NJ: Princeton University Press, 213–267.

Dilthey, Wilhelm (2010a). "The Origin of Our Belief in the Reality of the External World and Its Justification." In *Wilhelm Dilthey Selected Works II. Understanding the Human World*, edited by Rudolf Makkreel and Frithjof Rodi. Princeton, NJ: Princeton University Press, 8–57. (Original work published in 1890).

Dilthey, Wilhelm (2010b). "Contributions to the Study of Individuality." In *Wilhelm Dilthey Selected Works II. Understanding the Human World*, edited by Rudolf Makkreel and Frithjof Rodi. Princeton, NJ: Princeton University Press, 211–284. (Original work published in 1895–96).

Dilthey, Wilhelm (2010c). "Ideas Concerning a Descriptive and Analytic Psychology." In *Wilhelm Dilthey Selected Works II. Understanding the Human World*, edited by Rudolf Makkreel and Frithjof Rodi. Princeton, NJ: Princeton University Press, 115–210. (Original work published in 1894).

Dold, Malte and Rizzo, Mario J. (2023). "Menger's exact laws, the role of knowledge, and welfare economics." *The Review of Austrian Economics* 36: 163–182.

D'Oro, Giuseppina and Søren Overgaard (2017). *The Cambridge Companion to Philosophical Methodology*. Cambridge: Cambridge University Press

Ermarth, Michael (1978). *Wilhelm Dilthey: the critique of historical reason*. Chicago: University of Chicago Press.

Feher, István M. (2015). "Hermeneutics and Ontology." In *A Companion to Hermeneutics*, edited by Niall Keane and Chris Lawn. John Wiley & Sons. 162–171.

Forster, Michael and Kristin Gjesdal, eds. (2019). *The Cambridge Companion to Hermeneutics*. Cambridge, Cambridge University Press

Gadamer, Hans-Georg (2013). *Truth and Method*. Second revised edition, translation revised by Joel Weinsheimer and Donald G. Marshall. New York: Bloomsbury. (Original work published in 1960).

Gjesdal, Kristin (2014). "History and Historicity." In *The Routledge Companion to Hermeneutics*, edited by Jeff Malpas and Hans-Helmuth Gander, 299–309.

Gjesdal, Kristin (2017). "Hermeneutics and the Question of Method." In *The Cambridge Companion to Philosophical Methodology*, edited by Giuseppina D'Oro and Søren Overgaard, 337–355.

Gjesdal, Kristin (2019). "Hermeneutics and the Human Sciences." In *The Cambridge Companion to Hermeneutics*, edited by Michael Forster and Kristin Gjesdal, 354–380.

Haller, Markus (2004). "Mixing Economics and Ethics: Carl Menger vs Gustav Von Schmoller." *Social Science Information*, 43:1, 5–33.

Heidegger, Martin (1977). *Sein und Zeit. Vsk. 2. Gesamtausgabe*. Frankfurt am Main: Vittorio Klostermann (original published in 1927)

Holborn, Hajo (1950). "Wilhelm Dilthey and the Critique of Historical Reason." *Journal of the History of Ideas* 11/1, 93–118.

Kidron, Carol A. (2012). "Alterity and the Particular Limits of Universalism Comparing Jewish-Israeli Holocaust and Canadian-Cambodian Genocide Legacies", *Current* Anthropology, 53:6, 723–754.

Kinzel, Katherina (2018). "Inner Experience and Articulation: Wilhelm Dilthey's Foundational Project and the Charge of Psychologism."

Hopos: The Journal of the International Society for the History of Philosophy of Science 8/2, 347–375.

Kinzel, Katherina (2019). "The history of philosophy and the puzzles of life. Windelband and Dilthey on the ahistorical core of philosophical thinking." In *The Emergence of Relativism: German Thought from the Enlightenment to National Socialism* edited by Martin Kusch, Katherina Kinzel, Johannes Steizinger and Niels Wildschut, 26–42.

Kusch, Martin (2019). "Das Zeugnis der Holocaustüberlebenden: Gewissheit, Skeptizismus, Relativismus", *Deutsche Zeitschrift für Philosophie*, 67:6, 979–991.

Kusch, Martin, Katherina Kinzel, Johannes Steizinger and Niels Wildschut, eds. (2019). *The Emergence of Relativism: German Thought from the Enlightenment to National Socialism*. London: Routledge.

Lloyd, Geoffrey (2007). *Cognitive Variations. Reflections on the Unity and Diversity of the Human Mind*. Oxford: Clarendon Press.

Louzek, Marek (2011). "The Battle of Methods in Economics. The Classical Methodenstreit—Menger vs. Schmoller.' *American Journal of Economics and Sociology*. 70:2, 439–463.

Maclean, Michael J. (1982). "Johann Gustav Droysen and the Development of Historical Hermeneutics." *History and Theory*, 21:3, 347–365.

Makkreel, Rudolf A. (1975). *Dilthey. Philosopher of the Human Studies*. Princeton, NJ: Princeton University Press

Makkreel, Rudolf A. (1982). "Husserl, Dilthey and the Relation of the Life-World to History." *Research in Phenomenology* 12, 39–58.

Makkreel, Rudolf A. and Rodi, Frithjof (2019). "Preface to All Volumes" In *Wilhelm Dilthey: Selected Works, Volume VI*, edited by Rudolf A. Makkreel and Frithjof Rodi. Princeton NJ., Princeton University Press. vii–x.

Makkreel, Rudolf (2021), "Wilhelm Dilthey." In *The Stanford Encyclopedia of Philosophy* (Spring 2021 Edition), edited by Edward N. Zalta. URL = <https://plato.stanford.edu/archives/spr2021/entries/dilthey/>.

Malpas, Jeff and Hans-Helmuth Gander (2014). *The Routledge Companion to Hermeneutics*, edited by Jeff Malpas and Hans-Helmuth Gander. London: Routledge.

Mannheim, Karl (1952). "The Problem of Generations." In *Karl Mannheim: Essays*, edited by Paul Kecskemeti. London: Routledge, 276–322.

Nelson, Eric (2014). "Heidegger and Dilthey: Language, History, and Hermeneutics." In *Horizons of Authenticity in Phenomenology, Existentialism, and Moral Psychology*, edited by Megan Altman and Hans Pedersen. New York: Springer, 109-128.

Pinxten, Rik (1976). "Epistemic universals. A contribution to cognitive anthropology." In *Universalism versus Relativism in Language and Thought: Proceedings of a Colloquium on the Sapir-Whorf Hypotheses*, edited by Rik Pinxten. Berlin, Boston: De Gruyter Mouton, 117-176.

Plantinga, Theodore (1980). *Historical Understanding in the Thought of Wilhelm Dilthey*. Toronto: University of Toronto Press Scholarly Publishing.

Rickert, Heinrich (1962). *Science and History: A Critique of Positivist Epistemology*, edited by A. Goddard, translated by G. Reisman. Princeton, NJ: D. Van Nostrand. (Original published in 1902).

Rockmore, Tom (2003). "Dilthey and historical reason." *Revue internationale de philosophie*, 226:4, 477–494.

Scharff, Robert C. (2019). *Heidegger Becoming Phenomenological: Interpreting Husserl Through Dilthey, 1916–1925*. Lanham: Rowman and Littlefield.

Shalin, Dmitri N. (1986). "Romanticism and the Rise of Sociological Hermeneutics." *Social Research,* 53:1, 77–123

Troeltsch, Ernst (1946). *Gesammelte Schriften. Bd. 3, Der Historismus und seine Probleme; Buch 1, Das logische Problem der Geschichtsphilosophie*. Tübingen: Mohr.

White, Daniel and Wang, Alan (1995). "Universalism, humanism, and postmodernism." *American Psychologist*. 50, 392–393.

A Note to "Meaning in Time"

JAAKKO REINIKAINEN

1. Introduction

As the title suggests, this paper is something of a leftover – or perhaps a new branch – to my "Meaning in Time: on temporal externalism and Kripkenstein's skeptical challenge" (Reinikainen 2022). In that work I essayed to portray my understanding of the skeptical challenge uncovered by Saul Kripke's (1982) reading of Wittgenstein's later works in a nutshell as to its nature and solution. Here, my task is to dig a little deeper into the key phrase of the earlier paper, namely the claim that meanings, facts grounding meaning facts, and ascriptions of meaning have an important historical dimension to them. These specifications are due not only to existing thoughts I could not fit into the earlier paper, but also due to conversations I had the pleasure to exchange in the Philosophical Society of Finland 2022 colloquium in Oulu – especially with Teemu Tauriainen – to which I had the honor to contribute.[1]

At bottom, I believe that the skeptical challenge is best understood as logical as opposed to metaphysical in nature. Yet certain metaphysics of meaning are more compatible with its conclusions than others. By "metaphysics of meaning" I mean primarily the question of what is the *nature* of meaning, assuming already that there are such things as meanings. In particular, the main contestation of this paper is that the causal-historical account of reference originating from Kripke (1980) and as developed by Michael Devitt (1996) is well-suited to make sense of the somewhat esoteric-sounding expression "historicity of meaning". In sum, we can explain the historicity of meanings by the historicity of the referents, for

[1] The paper also benefitted, especially in its clarity of exposition, from the insightful comments made by two anonymous referees.

meaning is (by its nature) an aspect of causal-historical reference. (Put more conservatively, that is one important property of a word one can mean by "meaning".)

Initially I shall explicate the account on its "home-turf", or in the case of proper names. The more programmatic purport of the paper, however, is to expand the account from proper names to other term types, most importantly natural kind terms, and ultimately to all terms. Perhaps that is an overreach already in programmatic terms, but if so, the failure will be all the more fruitful.

The paper will proceed as follows. In the second section I shall briefly sketch the skeptical challenge as uncovered by Kripke, with focus paid on the claim that the challenge is best understood as "logical" as opposed to "metaphysical" in nature. Although I purport to make the discussion self-standing, I will make some use of the exposition launched in the earlier paper.

The third section expounds on certain key features of Devitt's (meta)semantic program as well as on the point of why it is naturally posed to explain the historicity of meaning, taken as one major conclusion of the skeptical challenge. The core here is Devitt's "shocking idea" that, given a broadly Fregean approach to meaning as well as semantic externalism, at least some meanings (i.e., modes of presentation of the referents) must be understood as causal-historical in nature. If true, this thought will naturally explain the historicity of meaning, I shall argue. In summary, the argument is this:

> (1.) Meanings (of proper names) are causal-historical modes of reference.
>
> (2.) Causal-historical modes of reference are temporal and exist in time.
>
> (3.) Hence, meanings (of proper names) are temporal and exist in time; they have a history.

Historicity of meaning in this sense allows us to resolve the problem of finitude by understanding the facts that determine meaning facts as themselves temporal and finite in nature. This lets us give up absolute determinacy of meaning, which is the pivot of the problem of finitude.

2. The Skeptical Challenge as a Logical Problem

This section will briefly elaborate on the skeptical challenge about meaning as discovered by Kripke (1982). The exposition draws from my (2022); how exactly so will not be discussed in detail, but I trust that the reader will see the resemblance.

In its "raw" form, i.e. the form in which Kripke chose to cast it (as a helpful guide to introductory classes on Wittgenstein!) the skeptical challenge asks us to explain in virtue of what facts is it the case that a subject, let's call him Jones, who has up to a given moment learned to do addition with numbers less than 57, when asked to calculate "58 + 67", should answer (given his previous training and intentions) "125" and not, say, "5". That is to ask, what determines that Jones has up until now been following the addition function with his use of "+" as opposed to a seemingly arbitrary "quaddition" ("$\oplus$") function, according to which

$x \oplus y = x + y$, if $x, y < 57$

$= 5$ otherwise. (Kripke 1982, 8–9.)

There are three separate though interlinked problems from which the skeptical challenge consists of, at least according to the received view in the literature. These are called: the problem of finitude, the problem of error, and the problem of normativity. For the aims of this paper, it suffices to focus on the problem of finitude, which in any case I believe to be the most important problem as I will argue in my oncoming doctoral dissertation.

The problem of finitude is intuitively graspable on the basis of the example. We can begin with the observation that, whatever fact it is that determines which function (if any) Jones is following, it must be a fact that is at least partially about Jones. Why is that? Because it is Jones's arithmetical behavior that is under discussion. Even if it turned out to be the case that Jones is an adder by virtue of a divine decree, this decree would have to be *about Jones* for it to determine his behavior with the "+" sign. The second key observation is that Jones, *ex hypothesi* a normal human, is a finite being. He is only capable of exhibiting finite mathematical behavior, or finite dispositions to mathematical behavior. For example, he is not capable of calculating with *very* large numbers.

It is important to be clear that the skeptical question concerns Jones's *actual* dispositional states and behavior, con-

trasting these items with the *logically possible* states in which he alternatively, and incompatibly, can be in. The point of the problem of finitude is that because the actual item must be finite, it will always remain compatible with an indefinite number of alternative, mutually incompatible possible states in the sense of realizing any of them. The possible states can be formally rendered as functions, though following Wittgenstein, Kripke often calls them "rules" or just "meanings". What Kripke calls the "straight" solution to the challenge must, among other things, explain in virtue of what fact Jones's mathematical behavior is governed by one such unique rule (1982, 66). Strictly speaking it need not turn out that Jones is in fact an adder for the skeptical challenge to be solved: it could turn out that he is a quadder instead, though this would be surprising. So long as there is a determinate fact which he is, the challenge will have been solved. Relatedly, the limit of 57 where addition is revealed to be quaddition is arbitrary and can be replaced with any number with which Jones has not *ex hypothesi* yet calculated with.

In my "Kripkenstein semanttista realismia vastaan" ["Kripkenstein against semantic realism"] (Reinikainen 2021) I argued that semantic dispositionalism – perhaps the most popular straight solution candidate to the challenge – fails on the problem of finitude. The argument in "Meaning in Time" (2022) aimed to be more encompassing in claiming that there simply is no straight solution to the problem of finitude; it is an insoluble paradox. However, I also argued that the challenge turns into a paradox only against a certain assumption, which is not necessary, about the "rules" or "meanings" among which the skeptic demands determination. In particular, it is implicit in Kripke's exposition that such rules or meanings must be "absolutely determined". While the literature knows a number of definitions that plausibly fill the role, I continue to find Alexander Miller's formulation as the most concise and helpful:

> In the case of a descriptive expression such as "+," whatever fact that is proposed as making it the case that "+" means the addition function must be inconsistent with the hypothesis that "+" means some other function, such as quaddition. In the generalized version of the argument, which applies to both descriptive and non-descriptive language, this becomes: whatever fact that is proposed as making it the case that rule R_a is the rule governing Smith's use of expression E must be inconsistent with the

> hypothesis that the rule governing his use of E is R_b, where R_a and R_b are such that for some possible use Δ of E, Δ is correct according to R_a but incorrect according to R_b. (Miller 2010, 460.)

The main lesson of the skeptical challenge is that we should reject absolute determinacy of meanings, facts determining meaning facts, and meaning ascriptions. What this means is that the semantic values of expressions (e.g., the addition function for "+", or the set of all tables for "table") cannot be individuated by rules that would govern all logically possible applications of the expression such that it would be determined for every potential application whether it was correct or incorrect according to the rule. Although there is nothing, I don't think, in the skeptical challenge that would forbid us from modeling semantic values theoretically (i.e. for the purposes of descriptive semantics relying on model theory) as if they were absolutely determined, insofar as these models are applied to expressions as used in actual natural languages, they will always remain indeterminate due to the fact that the semantic values of actual expressions must be determined diachronically, or temporally, and are nowhere absolutely finished.

What, then, does it mean for the meanings and reference of words to be "temporally determined"? While a full answer to that question will have to wait, here we can make some progress by looking at the idea in rough principle. In a somewhat extreme sense, temporal determination of meaning means that whether a given particular referent belongs to the extension of a given term is a temporal matter in the sense that it will have to be decided in time whether the referent belongs to the extension or not.

This might seem wildly implausible: surely it is a different issue whether an object "fits" a standard (of reference) and whether we can know that it does. While knowing whether the animal in the bushes is correctly called a "dog" is something that happens in time, and must be decided in that sense, the issue of whether "dog" correctly applies to the thing in the bush is independent of epistemic issues, assuming the usual meaning of "dog".

To get the point right, the indeterminacy with which the skeptical challenge trades does not (primarily) concern (i) the epistemic issue of whether and how we can know that a given object fits a given standard or not, nor (ii) that objects can

fit standards factually "by themselves", i.e. independently of our beliefs and knowledge, but rather (iii) the point that what standard (semantic value) is selected by a token expression in context cannot everywhere be decided by a further standard. The way a rule is to be applied in practice cannot everywhere be settled by further rules, for these too would then require rules for their application etc.

The logical nature of the problem of finitude means also that the arithmetical example is in no way special in kind. The same basic question can be raised wherever we have a standard of correctness of some kind, as I will next illustrate with an example. Consider that I am in the middle of purchasing some dry goods that I have yet to see for myself, and make the following statement while measuring a length in the air between my index finger and thumb:

> If it is *this* long, I will buy it.

Now, at the time of making the assertion, did I mean the distance as measured between the insides of my fingers or between the nails? (Assume for sake of argument that this comes to relevance later.) Plausibly, at the moment I did not explicitly intend one standard of length over the other. But is there still some other fact which might settle which length (or any number of other logically possible alternatives) was meant in the moment aside from my intentions? What kind of a *fact* could even in principle be suitable here? Psychological facts about humans, trade conventions, and other contextual matters might provide plausible answers. Ultimately it could nonetheless turn out that the issue is vague: there was no fact of the matter which length was the intended standard at the time of the utterance. Even if me and my trade partner later come to an agreement as to which length was meant, this as such does not mean that the matter was determined at the moment of utterance.

I think one reason why Kripke chose to use addition as the paradigmatic example in his presentation was to avoid mixing in the type of "mundane" or ordinary kind of indeterminacy such as vagueness that we encounter, e.g., in lay measurement. To apply the problem of finitude proper in this case, we would first need to think of the standard used (i.e., the intended length) as having the form of a rule, in other words, a logical form. (To simplify, we can think of the "logical form" along Miller's characterisation quoted above.) The logical form itself may contain vagueness of many sorts: the

rule may not specify whether the intended length is to be measured from inside my fingers or between the nails. However, the important point is that some possible lengths are intuitively excluded by this form containing vagueness. For example, it is not vague whether the intended length was the distance from my index finger to the tip of my nose, or to my toes, or to the Eiffel Tower. These lengths, we would say, are simply different standards governed by different rules, just like quaddition is a different standard from addition even though they share the same "input" (in the length case, my index finger). The problem of finitude arises precisely when we ask what facts exclude these alternative standards in the context of the utterance. The eventual point is to see that no "finite, temporal" fact can carry out this task.

That we don't ordinarily think that most possible standards *need* to be excluded to begin with to ensure smooth interaction is not an objection to the skeptic, which precisely goes to show that the problem is logical in nature. The skeptical challenge targets the assumption that our words and expressions, or even bodily gestures, in order to be meaningful, must select unique standards in the way of absolute determinacy, in which case it is always possible to raise the alternative possible standards. If the standard that is selected in context for the truth (or more generically "semantic correctness") of an utterance token must determine an infinite partition of correct-incorrect possible applications of the term while excluding an indefinite number of other partitions, there is no fact of the matter which standard is ever uniquely selected in a given actual context, simply because every actual context is finite.

This brief recapitulation was not meant to provide a foolproof argument, merely an illustration of the motivation for taking the problem of finitude seriously. In the fourth section, assuming that thoughts along these lines can be defended through various objections, I shall examine how Devitt's causal-historical account of reference can explain "historicity of meaning"; before that, the next section will provide a short introduction to Devitt's relevant ideas.

3. Devitt's Shocking Idea about Meaning

Behold:

> *The Shocking Idea*. The meanings of some words, including names and natural kind words, are causal modes of reference that are partly external to the head. (Devitt 2001, 477.)

Why is the idea "shocking", exactly? Or rather, whom is it likely to shock (or at any rate, mildly displease)? For one, it is shocking to anyone who thinks all meanings must be descriptive, or "in the head" of the speaker and her audience. It is shocking to anyone who thinks that meanings, as theoretical terms, must be ontologically distinct from empirical or observational terms. In the following I shall say a few words to alleviate these shocks, then go over Devitt's elaborations of his account as well as motivations for it.[2]

The first cause of shocked-ness is due to the familiar internalism-externalism debate in philosophy of language. Briefly, while internalists think that the referents of words are determined only by properties (or other items) internal to the speaker, externalists think that at least the referents of some words are determined by properties (or other items) that are external to the speaker. The causal-historical theory of reference advocated by Devitt identifies the external properties as causal-historical chains of "borrowing" and "grounding" that circulate in the speech community. It is worthwhile to point out, as Devitt does (1996, 162), that supposing one accepts a) externalism and b) the Fregean idea of meanings as modes of reference, one has no choice but to accept, *ceteris paribus*, the shocking idea. For if the meaning of a word is whatever determines reference, and if the referents of some words are determined by causal-historical chains of reference, then it follows that the meanings of some words are causal-historical chains of reference.

This note is too short to even summarize reasons for why one should (not) buy into either externalism or the Fregean

---

[2] There are other reasons to find the idea shocking that I will not touch on here. One such is the thought that meanings must offer normative reasons to use words one way rather than another, combined with the sentiment that merely empirical facts like causal-historical chains of reference are unfit to serve as genuine reasons for action or belief. Another is "direct reference theory", which states that proper names have no meaning at all, causal or descriptive.

approach to meaning, but I think that since both commitments are reasonable by themselves, that alone should take some edge off the shockingness of the shocking idea. How about the idea that, since meanings are (quite clearly) theoretical terms, and since the causal-historical chains are (at least in some sense) observable, we cannot identify the former with the latter? To be clear, I do not actually know of any objections to Devitt along these lines, but since I think it could be a natural remark to make, I want to give one reason how to deal with it.

To begin with, "meaning" is very clearly a theoretical term in the sense that we posit them in various theories for explanatory purposes. Some of the purposes include explaining observable effects like intentional behavior. However, this neutral observation as such does not forbid the identification of "meanings" in some observational vocabulary; in other words, there is no problem in "meaning" having both theoretical and observational uses. This would only be a problem under the assumption that there is an ontological difference between theoretical and observational terms such that the referents of theoretical terms exist in a different sense than the referents of observational terms do. But as Robert Brandom has argued, this is often not (if at all) the case. Rather, the difference between theoretical and observational terms is epistemic (or "methodological") in nature:

> Understood thus methodologically, the status of an object as theoretical or observable can change over time. When Pluto was first postulated, it was as a theoretical entity about which we could know only by making inferences from perturbations in the orbit of Neptune. With improvements in telescopy, looking at the calculated position of the hypothetical planetoid yielded the first observations of Pluto. It became, for the first time, observable. But *it* did not change ontological status; only its relation to us changed. Astronomers had been referring to the same planetoid, and knew things about it such as its orbit and mass, before it became observable—and would have done even if it had *never* become observable. A comparable story could be told about Mendelian genes. (Brandom 2015, 60.)

I see no *prima facie* reason why this thought could not be applied to meanings as causal-historical chains. Although the sense in which such chains are "observable" is a topic in and on itself, the fact that there is a persuasive reason to interpret

the theoretical-observational split non-ontologically is enough to shave some of the shockingness of the shocking idea, which claims that some meanings are "empirical" broadly speaking.

Now, onto Devitt's account about meanings as causal-historical chains in the case of proper names. To start off, the account as rendered here is mostly programmatic and thus scant in detail; in *Coming to Our Senses* (1996, 163) Devitt calls it "illustrative theory" (IT). But IT should suffice here to render the general idea clearly enough.

IT consists of three main parts which I will first roughly sketch, based mostly on Devitt (1996):

> *Grounding*. Following Kripke's lead, a proper name becomes first glued to its referent via an "initial baptism", which is the intentional act of using the name to refer to the (usually observationally present) object. Although baptism is a form of intentional action, the link that determines reference is causal and not descriptive in kind, although it is possible that some descriptive intentions are necessary for determinate baptism to occur at least in case of natural kind terms. One of Devitt's original ideas is the possibility of "multiple grounding": the first link in the chain of reference is not privileged in any way, but rather the grounding of a name should be understood as a continuous, prolonged process which may also ensue in reference-change.
>
> *Borrowing*. Once the name has been (multiply) grounded, it can be borrowed from speakers who have been in causal contact with the referent to those who have not been in such contact and borrowed further from those speakers. While borrowing also is a form of intentional action, the key point is that it does not depend on the speaker or the hearer to have in mind a description (e.g., a belief) which singles out the referent necessarily and/or sufficiently.
>
> *Mental processing of D-chains*. While the first part deals with the speakers' relation to the referents, and the second part with speakers' relations to each other, the third part deals with what goes on inside individual speakers' heads when they refer by proper names. In schematic terms, the "D-chains" (designation-chains) are stored in the mental system under different type-files that must meet at least four criteria. (i) the files must be able to distinguish between physically (e.g., phonetically) distinct to-

> kens of names; (ii) the files must be able to distinguish between physically identical yet referentially distinct tokens of names; (iii) the files must be able to distinguish between physically distinct yet referentially identical tokens of names; (iv) the files must be able to distinguish between physically identical and referentially identical tokens of names.[3]

It goes without saying that a lot more would have to be said to make IT a respectable philosophical (let alone empirical) theory about names and naming. However, my next task is to argue how IT is compatible with the rejection of absolute determinacy and endorsement of historicity of meaning discussed in the previous section. With that purpose in mind, I shall say a few more words about grounding specifically, but otherwise the development of IT will be left for future work.

4. Why the Causal-Historical Account Explains Historicity of Meaning

Bareboned, the thesis that meanings have history (the "historicity of meaning") means that present uses of a word type depend for their meaning on earlier uses. According to a (broadly) Fregean theory, one property of a word denoted by its "meaning" is its mode of presentation of the referent, and one theoretical job of meanings is to determine the referent of a word. Putting these together, earlier uses of a term partly determine the meaning (i.e., mode of presentation) of the term in the present by determining what referent the term has had in the past. For example, the reason why the past uses of the name "N.N." determine (in part) the meaning (and thereby the reference) of "N.N." in the present is that the present mode of presentation of N.N. by "N.N." depends on the past modes of presentation of N.N by "N.N.", which depend on past uses all the way down to the original use. And the

---

[3] The fourth criterion may seem strange, and it certainly is a rarity, but it is still a possibility which IT should be able to account for. Devitt provides as an example a Batman-style scenario where the person was known in both of his lifeworlds by the physically same name, say "Bruce", without this double-life being generally known (1996, 167). Evidence that subjects would still file these names differently is that they would intuitively behave differently in relation to tokens of "Bruce-AKA-Batman" and "Bruce-AKA-the-billionaire", even when these persons were in fact the same, supposing it was generally not known that they were.

reason why they so depend on is that the present mode of presentation of N.N. just is in relevant parts the same as the past mode of presentation, namely the name "N.N." and the network of D-chains that underlie it.

The important point, then, is this. Devitt's causal-historical account explains historicity of meaning (at least in the case of proper names) by the simple fact that later links in the D-chain of a word depend for their existence and nature on the earlier links. Since the network of D-chains forms the mode of presentation of the name's referent, and because the mode of presentation is one property which can be identified as a word's meaning, it follows that present tokenings of a name depend on earlier ones for their reference and meaning. Thus, the account is compatible with the historicity of meaning – in fact, vindicates it.

Is IT also compatible with the rejection of absolute determinacy? I see no fundamental problem in interpreting it this way. As we saw in section 2, the skeptical challenge is primarily a logical problem that has to do with exclusion of alternative semantic values for a given token expression. The type of expression as such is irrelevant; we could equally well pose the challenge to a given proper name, say "Kripke", and ask in virtue of what fact does "Kripke" mean (or refer to) *Kripke* and not *Kripnam*, where "Kripnam" means (refers to) the disjunctive set "Kripke or Putnam".[4] However, according to IT, what the referent of a name is is in principle an empirical question, not a logical problem. There is more to be said here, but the key point is that for IT, since meanings of proper names are at bottom empirical, the question whether e.g., "Kripke" means Kripke or Kripnam, while always logically available to be asked, is no more troublesome than the question whether "London" and "Londres" refer to the same city or not. The crucial point concerns the description of grounding uses and intentions. If IT can explain how grounding uses work without appealing to atemporal facts, like facts about truth conditions of the grounding intentions, then it will be able to avoid the problem of finitude, for then facts grounding meaning facts will not have the problematic logical form of rules.

This "solution" to the skeptical solution is fundamentally not "straight" in the sense that it would take the skeptical question at face value and then provide an answer to it. As I

[4] The example is from Colin McGinn (1984).

argued in my (2022) (mostly following Martin Kusch's lead, cf. Kusch 2006), there is no straight solution to the skeptical challenge because the skeptical challenge is at bottom a successful *reductio ad absurdum* argument against a certain philosophical picture of language and meaning, part of which includes commitment to absolute determinacy. Once we give up absolute determinacy, we thereby "solve" the skeptical challenge (or at least the problem of finitude) by granting one of its main points as correct; the point being, there is no real logical problem to begin with, only an empirical one.

Of course, assuming this solution is valid, nothing much has been solved yet as regards how names actually work. Overcoming the skeptical challenge is not the end of problems; it only makes it possible to see true problems clearly. For the rest of this note, my aim is to say a few words on the grounding part of IT and how it needs to be adjusted in view of historicity of meaning and rejection of absolute determinacy. What is first needed is a robust account of referential intentions, which play an important role in the grounding of names and other terms, not to mention in their borrowing. To this effect, in the next section I shall look at an interesting proposal by Mario Gómez Torrente.

5. Empirical and Logical Indeterminacy

A central lesson of the skeptical challenge that I have focused on is that the price for avoiding absolute *in*determinacy words (i.e., the conclusion that no word has no determinate meaning whatsoever) is to give up absolute determinacy of words. This is another way of saying that there is no straight solution to the skeptical challenge and the problem of finitude in particular.

The correct follow-up, then, is to embrace indeterminacy in how the meanings of our words are determined. This is, in essence, the strategy that Gómez-Torrente has advocated independently of the skeptical challenge:

> I seek to provide a strong anti-descriptivist consideration about reference fixing for names and demonstratives, based on cases of referential indeterminacy, that has not, to my knowledge, been exploited in the previous literature on the topic. I then develop an account of reference fixing for these expressions which is compatible with antidescriptivism and which embraces the idea, hinted at by Kripke and others, that the relevant reference fixing conventions need not provide necessary and sufficient

> conditions for reference, but only imprecise roughly sufficient conditions for reference and reference failure. (2019, vii.)

There are two interrelated kinds of "referential indeterminacy" that are important to Gómez-Torrente's overall argument: indeterminacy over the success or failure of a given referential act and indeterminacy over what is being referred to by a given referential act. (To specify, at least in my terminology a referential act "fails" in the sense that it refers to nothing, not that it would refer to something else than what the speaker intended; the latter could be called "misreference" for convenience.) Next, I shall briefly exemplify how this kind of indeterminacy fits well with the shocking idea and historicity of meaning.

First an example of indeterminate referential success or failure in case of a proper name, inherited from *Naming and Necessity*:

> If [. . .] the teacher uses the name 'George Smith '—a man by that name is actually his next door neighbor—and says that George Smith first squared the circle, does it follow from this that the students have a false belief about the teacher's neighbor? The teacher doesn't tell them that Smith is his neighbor, nor does he believe Smith first squared the circle. He isn't particularly trying to get any belief about the neighbor into the students' heads. He tries to inculcate the belief that there was a man who squared the circle, but not a belief about any particular man—he just pulls out the first name that occurs to him—as it happens, he uses his neighbor's name. It doesn't seem clear in that case that the students have a false belief about the neighbor, even though there is a causal chain going back to the neighbor. (Kripke 1980, 95–96.)

Citing this example, Gómez-Torrente agrees with Kripke('s assumed view) that whether or not the pupils' use "George Smith" fails or succeeds to refer to George Smith is indeterminate in view of a competent speaker's linguistic intuitions (Gómez-Torrente 2019, 73).

Second, an example of indeterminacy of what is being referred to in case of a proper name:

> For another example, turn to the above (historically inaccurate) story often attributed to Evans. In the story, when Marco Polo inherits the name "Madagascar" he has both the intention of referring to whatever the Arab or Malay sailors referred to and the

> intention of referring to the island now known as "Madagascar "; these conflict, for they lead to different objects. At this point, it seems as if neither of those two referential intentions overrides the other, and it is intuitively uncertain whether "Madagascar," as a matter of what the reference fixing conventions determine, refers to either thing in the mouth of Marco Polo. (Of course, there must be a later time in the history of the transmission of the name (according to the story) in which "Madagascar" begins to refer to the island in the mouth of normal speakers.) (Gómez-Torrente 2019, 74.)

The comfort brought by the rejection of absolute determinacy is that we need not be philosophically uneased by these results. The empirical world is rife with indeterminacy, and naming is simply a part of the empirical world. *There is no logical reason why these matters would necessarily have to be decided one way or another*. That is one of the lessons of the skeptical challenge.

There is a ready objection to be made here. It is more plausible to grant that singular acts of reference by a proper name may be indeterminate at the intuitive level, but how could the very *grounding*, or the matter of bifurcation, of a name be indeterminate? Well, in the majority of cases where the referent is an individual there is no relevant indeterminacy present at the intuitive level; every competent speaker agrees who is the referent of "Saul Kripke".[5] The key point is that the kind of indeterminacy showcased in the examples presented here is "empirical" in a broad sense, not "logical" in the sense of the skeptical challenge. It is in the implicit parameters of the examples that *most* logically possible alternatives are excluded at the intuitive level. To think that behind the intuitions and the causal chains there must be a logically unique solution, formulable in the manner of Miller's scheme, is to succumb to the idea of absolute determinacy. Although in most cases the referent of a proper name referring to an individual will not be indeterminate or experience bifurcations, the empirical possibility is always there due to the simple complexity of the causal world. But there is no *further* worry about indeterminacy in the sense of the skeptical challenge because meaning (at least in the case of proper names) is at bottom empirical, not logical, in nature.

---

[5] But cf. the examples in Gómez-Torrente (2019, 75–76).

6. Conclusions

It is hard to say which is a more complicated phenomenon: the human linguistic classificatory system or the world which it tries to classify. Arguably what is most complicated of all is the combination of the two in reality. From this perspective it is somewhat surprising to find that a common descriptivist objection to the causal-historical account of reference is that it leaves the reference of many terms (most importantly natural kind terms) too indeterminate. In contrast, Gómez-Torrente (2019, 45). argues that a big problem for descriptivist theories of reference determination is precisely their aim to give necessary and sufficient conditions for determination, against the evidence of intuitions.

So, if we understand referential indeterminacy as a feature of the phenomenon under study as opposed to a bug of the causal-historical theory, then it will turn out that Devitt's shocking idea is a much more palatable proposal for understanding the nature of meaning even in case of natural kind terms and many others. In fact, the shocking idea works as a partial explanation for why kind-term reference is ridden with indeterminacy: because the causal D-chains themselves are so complicated.

This note sought to expand on some of the themes of its parent paper, most importantly what to make of the esoteric-sounding phrase "historicity of meaning". A demystifying virtue of the causal-historical account is that it is able, I believe, to explain the historicity of meanings by the historicity of the referents, for meaning is an aspect of causal-historical reference.

Tampere University

References

Brandom, Robert (2015). *From Empiricism to Expressivism.* Cambridge MA: Harvard University Press.

Devitt, Michael (2001). "A Shocking Idea about Meaning." *Revue Internationale de Philosophie*, 4, no. 218, 471–494.

Devitt, Michael (1996). *Coming to Our Senses: a naturalistic program for semantic localism*. Cambridge MA: Cambridge University Press.

Gómez Torrente, Mario (2019). *Roads to Reference: an essay on reference fixing in natural language*. UK: Oxford University Press.

Kripke, Saul (1982). *Wittgenstein on Rules and Private Language: An Elementary Exposition*. Cambridge MA: Harvard University Press.

Kripke, Saul (1980). *Naming and Necessity.* Cambridge MA: Harvard University Press.

Kusch, Martin (2006). *A sceptical guide to meaning and rules defending Kripke's Wittgenstein*. Montreal: McGill-Queen's University Press.

McGinn, Colin (1984). *Wittgenstein on meaning: An interpretation and evaluation*. New Jersey: Basil Blackwell.

Miller, Alexander (2010). "Rule-following skepticism." In *The Routledge companion to epistemology*, edited by S. Bernecker & D. Pritchard, 454-463. New York: Routledge..

Reinikainen, Jaakko (2022). "Meaning in time: on temporal externalism and Kripkenstein's skeptical challenge". *Synthese*, 200 (288), 1–27.

Reinikainen, Jaakko (2021). "Kripkenstein semanttista realismia vastaan." ["Kripkenstein against semantic realism."] *Ajatus*, 78 (11), 187–219.

Temporal Nature of Philosophy and the Concept of Duration in the Philosophical Method of Henri Bergson

KATARIINA LIPSANEN

This article explores the temporal nature of philosophy and the concept of *duration (durée)* in Henri Bergson's (1859–1941) philosophical methodology. The aim is to examine how time, particularly the concept of duration, is present in Bergson's philosophical approach and his understanding of the nature of philosophy itself. The analysis primarily relies on Bergson's works, including *Creative Mind* (1934), *Mind-Energy* (1920), and his 1916 speech delivered at the student residence in Madrid[1], while utilizing the definition of duration found in *Time and Free Will* and *Duration and Simultaneity*. The focus is on the role of duration in Bergson's methodology and the resulting implications for the nature of philosophy. I will exemplify the practical application of the concept of duration with examples especially from *Creative Evolution*.

Although there have been limited studies on the role of duration in Bergson's philosophical methodology, there exist works that explore duration and its relationship with *intuition*, which Bergson defines as his method. This distinction between intuition and methodology is necessary here, since in Bergson's writings intuition has also a strong association to a type of intellectual capability comparable to intelligence.[2]

---

[1] Cf. "Discours prononcé à la résidence des étudiants" in *Mélanges* (1972).

[2] In my perspective, the distinction between the two is not definitively clear-cut, as intuition, in fact, encompasses both aspects. However, due to the diverse nature of the concept, it is essential to acknowledge that a study solely focused on intuition does not automatically equate to an in-depth exploration of Bergson's methodology. More on this distinction cf. (Lipsanen 2021).

And while duration and intuition have been much studied attempts to reconstruct Bergson's method are scarce. This article is part of an effort to create this type of structured model of the Bergsonian method to make it available for current philosophical study. Very helpful for my study have been the studies by Jean-Louis Vieillard-Baron and Leonard Lawlor (2010) on intuition and duration but also studies by David Lapoujade (2018) and Dimitri Tellier (2008) that have provided support for my methodological interpretations.

Concept of duration and spatialized time

The nature of time holds a significant position in Bergson's philosophy, a subject he already delves into in his first doctoral dissertation, *Time and Free Will*. Throughout this work, Bergson consistently revisits the notion that the root cause of numerous classical philosophical dilemmas lies in the conflation of quality and quantity. One example he explores in *Time and Free Will* is Zenon's paradox of Achilles and the tortoise. Bergson contends that this paradox stems from a confusion between motion and space, which, by extension, leads to a flawed comprehension of the essence of time:

> Why does Achilles outstrip the tortoise? Because each of Achilles' steps and each of the tortoise's steps are indivisible acts on so far as they are motions, and are different magnitudes in so for as they are space.[…] This is what Zeno leaves out of account when he reconstructs the movement of Achilles according to the same law as the movement of tortoise forgetting that space alone can be divided and put together in any way we like, and thus confusing space with motion. (Bergson 2001, 113–114.)

This confusion, specifically, enables the paradox to arise. Instead of perceiving movement as an indivisible act, it is reduced to space and reconstructed through uniform immobilities. According to Bergson, motion itself lacks homogeneity, and any appearance of homogeneity is only found in the space it traverses (Bergson 2001 115, 140–221). This tendency to equate quality with quantity, as Bergson argues, is prevalent in our everyday lives and is inherent in the very nature of human understanding. Bergson observes how sensations such as pain, for instance, possess an intensi-

ty that can only be quantified by external factors unrelated to the pain itself. Our consciousness perceives pain as greater if it affects a larger portion of our body or if it is accompanied by additional sensations like nausea or heavy breathing. It is through the number of body parts affected or the additional sensations accompanying the pain that our consciousness grasps pain as a quantity. According to Bergson, pain in itself is a quality (Bergson 2001, 35–38).

For Bergson this indicates that the human mind has a tendency quantify qualities, often resulting in the complete neglect or initial unawareness of the original quality. The human mind brushes the immediate impression to the side and an intellectualized impression takes precedence (Bergson 2001, 90). This happens because the human mind tends to intellectualize and with our sensations, this happens almost instantly. In this process, quality is transformed into a form that the human mind can measure. (Bergson 2001, 39, 42–43, 48–49.) In *Creative Evolution*, Bergson argues that this tendency to intellectualize stems from the evolution of human intelligence. Our minds are oriented towards tool fabrication, and the spatialization and quantification of qualities are merely extensions of this inherent habit (Bergson 2001, 139). Once this spatialization or quantification has been done, the qualities become our tools.

The very same tendency of the human mind extends not only to our everyday impressions but also to more complex concepts, including the notion of time. According to Bergson, what we commonly refer to as time is fundamentally no different from our concept of space. Rather than understanding time on its own terms, it is comprehended solely in terms of space (Bergson 2001, 113–114, 181–183; 2007, 4). This results in a spatialized conception of time. But what does "spatialized time" mean? It implies that time is perceived as something measurable, a homogeneous continuum that can be divided into units and represented as a timeline(Bergson 1965, 57; 2001, 98). Bergson's central concern with this understanding of time is that it presupposes that the time under discussion has already passed, at least theoretically. Even when discussing future events, if time is depicted as a timeline, this representation implies that time has already been determined—it already has an endpoint (Bergson 2007, 2–3).

This presumed givenness of time, which forms the basis of our understanding, further influences the philosophical questions we attempt to address.

In *Time and Free Will*, Bergson argues that the debate on free will, encompassing perspectives such as determinism and indeterminism, is rooted in a flawed spatial understanding of time. This spatial conception reduces the question of free will to a mere inquiry into whether a choice between predetermined options was genuinely a choice. This limited perspective eliminates the possibility of unexpectedness or the emergence of entirely new solutions (Bergson 2007, 8.) After these criticisms, Bergson's own solution might seem rather unsatisfactory. He asserts that "freedom is the relation of the concrete self to the act it performs," yet he acknowledges that this relation is ultimately "indefinable" (Bergson 2001, 219). According to Bergson, when we analyze our supposed free actions, we inadvertently convert their duration into extensity. The indefinability of the relation between the concrete self and the act arises precisely because of our freedom. Bergson argues that any attempt to provide a positive definition of freedom would ultimately result in the triumph of determinism (Bergson 2001, 220). He posits that a truly free act is one that cannot be predicted, and any effort to analyze the conditions and antecedents of an act, as well as their connection to the act itself, inherently disregards the continuous flow of time. Such analysis treats both time and the act as events that have already transpired.

In short in *Time and Free Will* Bergson finds that human mind has a tendency spatialize, leading to conceptions such as determinism and indeterminism that rely on a spatialized notion of time. This general and the common conception of time can be useful in many instances, even necessary to many fields of science (Bergson 1965, 56–57; 2007, 3–4). However, Bergson argues that this conception is inadequate for philosophy and its aims, as it fails to capture the true essence of time (Bergson 1965, 65). Spatialized time, according to Bergson, is a reconstruction formulated by human intelligence, transforming time into a model of space for instrumental purposes (Bergson 1965, 57). Human comprehension, as well as science itself, necessitates the use of time as a tool. The purpose of science is to present the world in a manner that

enables us to act upon it. It predicts and measures, so we can best utilize it (Bergson 2007, 25–26).

While science relies on time as a tool, Bergson argues that philosophy is concerned with exploring the true essence of time itself.[3] Bergson's concept of *duration* presents a conceptualization of non-spatialized non-intellectualized real time. According to Bergson, our conceptions should be founded on our immediate experience (cf. Bergson 2001, 126–128). Bergson describes duration as the lived experience of time within ourselves, likening it to a flowing melody (Bergson 1965, 44; 2001, 100–101). It represents an ongoing continuity, characterized by constant movement and change, where individual moments can only be artificially separated. Just as a melody cannot be fully grasped by isolated notes, duration is understood by experiencing moments in a continuous succession (Bergson 1965, 49, 52). Real time, or duration, cannot be reduced to discrete units; it is an uninterrupted and heterogeneous flow.

The key differences between duration and spatialized time in their nature. Duration is a continuous and heterogeneous movement of change, while spatialized time is a homogeneous line that is considered already given, capable of being divided into measurable parts (Bergson 1965, 49). Real time is grounded in our immediate experience, whereas spatialized time is a reconstruction fabricated by human intelligence.

Duration has then an evident connection to the human mind. In Bergson's own words, duration no doubt unites or fuses into the continuity of our inner life (Bergson 1965, 44). Our immediate experience of time is rooted in the very continuity of our consciousness – without consciousness, there is no duration (Bergson 1965, 48). Memory is intrinsically linked to duration, as our perception of continuity and change relies on our ability to distinguish the past from the present and establish connections between them (cf. Bergson 1965, 44, 48–49). Without memory, our conscience would be in a sense reborn every moment, completely ignorant of the past. And there would be nothing to us except the present (Bergson 1965, 48; 2007, 137–138). These facts may appear self-evident,

[3] It should be noted that for Bergson philosophy is not a science and that these fields of knowledge are different by nature.

but they elucidate the central role of consciousness in understanding the concept of duration. Our mind or spirit (esprit) is, by nature, duration – indivisible continuity of change – and it is the duration that is most familiar to us.

But what relevance does this have for the nature of philosophy and Bergson's philosophical method? In his work *Creative Mind*, Bergson expresses, "These conclusions on the subject of duration were, as it seemed to me, decisive. Step by step they led me to raise intuition to the level of a philosophical method" (Bergson 2007, 18). Here, Bergson refers to the conclusions he reached while working on his doctoral dissertation. He was taken aback by the supposed conceptions of time found in the philosophical systems of sciences and some positivist thinkers, particularly Herbert Spencer in this case.

In his studies and later during his doctorate, Bergson was interested especially in Spencer's evolution, creativity, and progress-oriented thinking, but also the manner of his study. For Bergson, Spencer seemed to base his philosophy more directly on impressions of things and follow the facts more closely than any other philosopher. (Bergson 2007, 2; Verdeau 2007, 364–366.) One of the central ideas in Spencer's work, *The First Principles* (1862), was to observe how the universe becomes increasingly heterogeneous and differentiated from a previously homogeneous state, placing evolution at the core of Spencer's philosophy (Weinstein 2019). Bergson believed that *The First Principles* fell short in its understanding of mechanics, and he desired to follow Spencer's path, almost redoing Spencer's work with special attention to this particular detail. However, as he began this work, he noted the problems of spatialized time as demonstrated above: it discards the time itself and measures something else, namely, the spatialized representation of time. (Bergson 2007, 2–3.)

The search for the real time led Bergson to develop this new method that he termed intuition. Interestingly, the discovery of duration and intuition occurred almost simultaneously for Bergson, even though he fully articulated his own conception of intuition much later.[4] Based on Bergson's de-

[4] This is mentioned for example by Gilles Deleuze (Deleuze 1988, 13). Despite this, in *Time and Free Will* Bergson mentions intuition at least once

scriptions of intuition, a reference to a direct immediate vision in the context of duration already suggests that some idea of intuition was already present early on (Bergson 2007, 2–3). Bergson tries to dissociate his conception of intuition from the history of the concept, mentioning Schelling and Schopenhauer by name (Bergson 2007, 18). He emphasizes that the object of intuition is a primary reason for differentiating his concept from others, as intuition is often historically associated with grasping eternal ideas or principles. Bergson describes the initial use of his method as follows:

> I had chosen first of all to try out my method on the problem of liberty. In so doing I should be getting back into the flow of the inner life, of which philosophy seemed to me too often to retain only the hardened outer shell. (Bergson 2007, 15.)

For Bergson, *Time and Free Will* was the first attempt at reanalyzing some of the classical problems of philosophy with his newly discovered conception of duration.

Bergson's methodology

Now, shifting our focus to Bergson's methodology itself, I will begin by providing a general overview and then delve into the role of duration within this process. Based on my analysis, Bergson's philosophical method can be seen as comprising two distinct moments. The first moment is scientific in nature, while the second, as stated by Bergson himself, is truly philosophical (Bergson 1972, 1197).

In the first moment of his methodology, the philosopher immerses himself in the latest research conducted in the scientific field relevant to the topic at hand (Bergson 1972, 1197). For example, if we are interested in the nature of life, we should familiarize ourselves with evolutionary biology. If, on the other hand, we are interested in conscience or memory, we should focus on psychology or neurology, for example. This demonstrates Bergson's aspiration to adhere to scientific facts as closely as possible in a Spencerian manner.

---

in a way that clearly illustrates the effort to grasp our immediate impressions (Bergson 2001, 114).

In Bergson's view, science and philosophy have different goals. Science does not fully grasp what philosophy aims for. Science, or intelligence, is focused on measurement, spatialization, and the creation of typologies and categories (Bergson 1965, 48). It aims to intellectualize its objects and make them useful and accessible for future actions. We are not dealing with reality as such, but rather with the version of reality that our human intelligence can grasp. According to Bergson, philosophy, on the other hand, is interested in the essence and nature of objects as such. (Bergson 2007, 18–19; 102–103.) It is then clear that simply taking facts from the scientists is not enough.

Our philosophical inquiry commences by delving into the pertinent scientific disciplines and thoroughly examining their research, acquiring proper directions for the study. Although Bergson's explanations regarding this process are limited, his work *Mind-Energy* suggests that science does not merely provide us with these guidelines; instead, it falls upon the philosopher to identify and interpret them. Bergson also implies that this process specifically involves recognizing inclinations present within scientific studies. In *Mind-Energy*, Bergson exemplifies the formulation of these guidelines through his lecture on the relationship between life and consciousness, as well as the function of consciousness in living beings. As consciousness operates through the brain in human beings, Bergson initiates his analysis by examining the function of the brain within the nervous system. He emphasizes that while certain reactions to external stimuli involve the brain, there are numerous instances where the nervous impulse bypasses the brain and directly travels to the spinal cord. Bergson contends that the brain's involvement arises when a choice is required rather than an automatic response. By considering this and various other facts, we arrive at the conclusion that the brain is an organ of choice (Bergson 1920, 11–12). Ultimately, Bergson argues that consciousness embodies freedom and creativity in the evolutionary process of life, and the brain, as an organ of choice, serves as an early indication of this discovery.

The directions derived from scientific studies do not lead us directly to the truth. They provide us with guidance and point us in the right direction, but alone they are insufficient.

Bergson appears to view coherence or convergence of facts as the criterion for determining the correctness of the direction.

> In short, we possess even now a certain number of *lines of facts* which do not go as far as we want, but which we can prolong hypothetically. […] Each, taken apart, will lead us only to a conclusion which is simply probable; but taking them all together, they will, by their convergence, bring before us such an accumulation of probabilities that we shall feel on the road to certitude. (Bergson 1920, 4, emphasis added.)

Philosophical knowledge and its methodological process rely on coherence and the gradual accumulation of probability for their validity and credibility. The outcomes of philosophical investigations are inherently speculative, and it is not possible to assert absolute certainty in our studies. Bergson does not specify the reason for the speculative nature, but it can be assumed that it arises from two factors. Firstly, the progress of science provides valuable information that can be drawn upon for philosophical studies. Secondly, the utilization of intuition and the application of duration to elaborate on scientific knowledge are not without uncertainty. Bergson asserts that philosophy, similar to the sciences, should be progressive and self-correcting, with subsequent philosophers building upon the work of their predecessors and incorporating new information they have acquired. (Bergson 1920, 1–4.)

Once we have acquired insights from the sciences regarding our subject of interest and outlined the general framework of our study, we must address the remaining gap, as the first moment alone does not lead us directly to the truth. It is in the second moment of the method, the distinctively philosophical phase, that duration and intuition come into play. Regarding duration, its significance in the philosophical method can be illustrated through the following two quotes from Bergson:

> [L]et us in a word become accustomed to see all things *sub specie durationis* […] (Bergson 2007, 106).

> In fact, the more we accustom ourselves to think and to perceive all things *sub specie durationis*, the more we plunge into real duration (Bergson 2007, 132).

Sub specie durationis, "from the perspective of duration" refers to Spinoza's phrase *sub specie aeternitatis*, "from the perspective of eternity". The purpose of this contrast is clear: philosophy should move from the eternal, constant, and unchangeable perspective to the perspective of change (Bergson 2007, 6, 18–20). Bergson proposes that instead of seeking eternal ultimate principles, laws, or ideas, philosophers should direct their attention towards everything that changes and the inherent changeability of all things. The task of philosophy is to study reality from the perspective of duration, or rather study reality as change or duration.

In this very perspective, philosophy can be seen as the counterpart to science , which, according to Bergson, studies reality as static and homogeneous, aspiring to measure and spatialize it (Bergson 2007, 102–103). This also reveals the fundamental elements of Bergson's metaphysics, where reality consists of two basic factors: matter and spirit. Spirit represents the ever-changing and creative aspect of reality—it is inherently duration or the source of all duration. In Bergson's philosophy, everything in reality is formed of these two forces. The division of labor between science and philosophy specializes in each of these basic forces: philosophy focuses on spirit, while science focuses on matter (Bergson 1972, 887; 2007, 24–25). Ideally, the knowledge from each field complements the other, leading to a progressively more complete understanding of reality (Bergson 1965, 5; 1920, 7).

Before going further into philosophy's perspective, it is relevant to clarify the perspectives of science. What is the relationship between matter, space, and spatialization in the realm of science? According to Bergson, scientific thought, or intelligence, is modeled after matter; it treats everything as if it were matter. This perspective gives rise to spatialization. One might wonder if science is also concerned with change. Indeed, it is, but not in a similar way as philosophy. Bergson argues that sciences like biology do study change, such as evolution. However, this study involves transforming change into inertia and halting the movement of change. Science creates stages or phases to represent change in an analyzable form. (Bergson 1911, 170–172, 206–207.)

To use Bergson's analogy of a melody, grasping the change of melody as change (philosophy) means experiencing the

melody as it unfolds. On the other hand, analyzing the melody, treating it as if it were composed of individual phases strung together and effectively arresting the original movement of change (science), involves breaking it into notes and creating a notation. After making this distinction, one might be inclined to conclude that, for Bergson, science deals only with intellectualized constructions that lack a true connection to reality. However, I interpret Bergson's view as suggesting that science simply approaches reality from the perspective of matter and that it undeniably acquires knowledge that is no less valuable than the knowledge gained by philosophy. Notation provides us with a tool to analyze harmonies and compare melodies, even if it may lose sight of duration. This knowledge, while different from philosophical knowledge, is still valuable and true.

The task of philosophy is to study reality from the perspective of change: spirit as duration. According to Bergson, as we already previously mentioned, human intelligence tends to reconstruct everything changing spatially, time as a timeline as if it were like matter, measurable, and extensive object. In other words, the scientific outlook is the natural perspective of the human mind. The philosophical perspective on the contrary goes against human nature – human beings are not evolved to perceive change as change (Bergson 2007, 61–62). Instead, it is natural for the human intellect to freeze change into a particular form and view it as static. Time is not studied as it endures, but rather as a frozen timeline. Evolutionary changes are not observed as they occur, but rather as freeze-frame images, represented by different frozen phases (Bergson 1965, 60).

Duration and intuition

The question is, therefore, how despite this philosophy is capable of regarding change as change, if it is against human nature. This is possible with intuition. Bergson defines intuition in the following manner:

> The intuition we refer to then bears above all upon internal duration. It grasps a succession which is not juxtaposition, a growth from within, the uninterrupted prolongation of the past into a present which is already blending into the future. It is the

> direct vision of the spirit by the spirit,—nothing intervening […]. (Bergson 2007, 30, translation altered.)

Intuition finds its foundation in the inner duration inherent within human beings themselves (Bergson 2007, 20). As mentioned earlier, duration is intimately connected to the human mind and its nature; in fact, the human mind and consciousness are inherently characterized by duration. Furthermore, we have discussed how the true nature of time as duration can be directly apprehended through inner experience. Within ourselves, we can directly perceive the flow of time and the ever-changing nature of our inner states.

Bergson's methodology rests on a Cartesian premise, wherein one duration that we can know with absolute certainty is intricately intertwined with our own existence—the duration of our own being.[5] This is the key role of duration in his methodology. The question that arises is how we can utilize this immediate grasp of our own duration. In his methodology, Bergson seeks to extend the certainty and immediacy of our experience of duration to other objects of our knowledge, including things that exist external to ourselves (Bergson 2007, 20). However, the challenge lies in how we can establish a connection between the certainty and directness of our knowledge of inner states and the knowledge of other objects. There is an evident connection between our durations and other beings:

> How do we pass from this inner time to the time of things? We perceive the physical world and this perception appears, rightly or wrongly, to be inside and outside us at one and the same time […]. To each moment of our inner life there thus corresponds a moment of our body and of all environing matter that is "simultaneous" with it; this matter then seems to participate in our conscious duration. (Bergson 1965, 45.)

According to Leonard Lawlor, there exists an "infinity of other possible durations in my self" because my duration is a part of the greater "whole of duration", which suggests our participation in a universal duration (Lawlor 2010, 33–34). There are many similar interpretations with different empha-

[5] Similar interpretations have been made for example by Jean-Louis Vieillard-Baron (cf. 2004, 63).

ses. Michel Weber, for instance, explicitly asserts that our duration allows us to connect with the uninterrupted continuum of durations (Weber 2005, 129).

But to give this idea a more methodologically clear form, I introduce here Bergson's idea of reasoning by analogy.[6] In order to extend intuition to other objects, we must perceive the object of our study as analogous to our inner duration. Just as everything in reality comprises both spirit and matter, the object of study, as something that undergoes change, must also include spirit, which inherently possesses duration. Lapoujade and Tellier highlight that Bergson's analogy is not between two identical things, but between two entities that share something in common (Lapoujade 2018, 45; Tellier 2008, 425). The object of our study and our inner continuity has a shared nature and that is duration (cf. Bergson 1965, 45; Tellier 2008, 425–426). We can perceive objects in reality as resembling duration-like change because we ourselves embody the very same temporal change. An even simpler way to state this would be to say that we can look at reality as a temporal changing thing because we are ourselves temporal changing beings.

In addition to analogy, the concept of sympathy is also relevant here. Our shared nature with reality and other beings enables us to intellectually sympathize with the object of our study. There is not only something similar between us, but something that is the same: duration. We can sympathize with this duration, which allows us to understand other objects "from within" (Lapoujade 2018, 40). In this sense, sympathy refers to the same thing as the Bergsonian analogy. Sympathy seems to be a less technical term, but it demonstrates the act by which our understanding of the nature of the object occurs, which is not purely analytical or intellectual. As Jean-Louis Vieillard-Baron puts it, the intuition of duration is not the same as an idea of duration (Vieillard-Baron 2004, 50–52). While the immediate experience of our inner

[6] David Lapoujade has made a detailed study of this type of reasoning in Bergson's philosophy, which I have adopted here (Lapoujade 2018, 44–52). Bergson mentions this type of reasoning in a couple of cases, but does not elaborate much on this idea (Bergson 1965, 46; 1992, 438; 1911, 270).

duration is certain, extending it to other objects necessarily involves speculation. (cf. Bergson 1920, 4; 1965, 46.)

Finally, what are the results of Bergson's methodology? How does the concept of duration contribute to a philosophical study? I will use Bergson's work *Creative Evolution* here as an example. In said work Bergson studies the nature of life as an evolution. His view is contrasted in the beginning with views such as mechanism and finalism. Without going to detail, both of these views essentially argue that evolution is predetermined in some aspect: either the end goal (finalism) is given or each of the changes of evolution are connected by necessity; one change is necessarily followed by another and there are no other possibilities. According to Bergson true change and duration is excluded from these views and wishes to introduce a point of view that regards evolution as truly change.

Bergson analyzes various evolutionary lines and the tendencies that manifest within them. As Bergson regards the movement of life and the changes that occur in it, he draws a parallel to our own inner experience of change. Inner duration and change are understood as analogous to the changes and duration of life. This analogy is stated very directly: "Such is my inner life, and such also is life in general" (Bergson 1911, 272). Just as we cannot distinguish specific states in our inner experience of duration, such as states of mind and the transitions between them (except artificially), we also cannot do so with the movement of life or the changes of evolution. We can acknowledge that we are a different person than we were 10 years ago and point out certain differences in our personality. Similarly, we can identify variations in the characteristics of a species from thousands of years ago to the present. However, these differences are labeled using static states of "then" and "now". The actual ongoing change can only be grasped by looking at our inner experience of change and the continuous duration of time that we feel within ourselves in every moment. Through this lens, we can think and immediately apprehend the change in the movement of life, almost as if we are experiencing it within ourselves.

Bergson arrives to the idea of *vital impetus* (*élan vital*) according to which, the movement of life is created by the interplay of two opposite forces spirit and matter. Spirit

signifies the effort to create as freely as possible whereas matter provides the means of creation, but with limited possibilities. Bergson's ultimate idea is that life's aim is to overcome the limits of matter and create as freely as possible. He arrives to this conclusion especially by studying the evolutionary path to man. As previously mentioned, Bergson has noted that human consciousness has something to do with choice and freedom. Bergson argues that this development of human conscience demonstrates the life's effort create more freely. Development of human intelligence that is modeled after matter, demonstrate the life's effort to overcome limitations of matter. Human cognition has triumphed over these limitations by acquiring tools to manipulate matter. Through the acquisition of tools—both physical and intellectual—as well as the development of language, human beings have attained the capacity create ever more freely. (Bergson 1911, 278–280.)

Drawing on the insights of evolutionary biology, tendencies toward this inherent need for creativity can be discerned in various lines of evolution. However, it is through our understanding of our own consciousness and its duration that the principle of vital impetus becomes most evident. Similar to our individual duration, the duration of life itself is finite—it undergoes maturation and different lines of evolution become specialized. Bergson's argument suggests that these specializations are directed towards discovering the optimal means of surmounting the limitations imposed by matter. Thus, it is through the analogy and comparison of our inner duration with life's duration that Bergson identifies the nature of life as a vital impetus—an incessant striving to create with increasing freedom.

Bringing duration to the objects of our study completes our understanding of their nature. It goes beyond merely providing a more comprehensive definition or image. With the inclusion of duration and intuition, the objects of our contemplation come alive, so to speak:

> Thanks to philosophy, all things acquire depth,—more than depth, something like a fourth dimension which permits anterior perceptions to remain bound up with present perceptions, and the immediate future itself to become partly outlined in the present. Reality no longer appears then in the static state, in its

> manner of being; it affirms itself dynamically, in the continuity and variability of its tendency. (Bergson 2007, 131.)

For Bergson philosophy animates the objects of its study. It does not view things as immutable or eternal, but rather acknowledges the inherent temporality within them. Philosophy regards reality as temporal, and ever-changing and preserves the temporal nature of the objects of its study. This perspective also has profound implications for philosophy itself, as philosophical thought becomes inherently temporal as philosophical thought itself is thinking in duration.

Conclusion

In conclusion, Bergson's method of intuition is grounded in the notion that philosophy seeks to comprehend reality through one of its fundamental elements: spirit, duration, or change. Together with science, which examines reality from the perspective of matter, they complement each other and form a more comprehensive understanding of reality. Bergson's concept of duration offers philosophy a unique approach to its inquiry.

Within the framework of Bergson's method, duration assumes a significant role in the second phase, namely intuition. It follows the scientific phase, which provides philosophy with general orientations and a framework for the philosophical intuitive approach. The knowledge acquired through familiarization with scientific discoveries, guided by intuition, can be enriched in the second phase, where duration plays a crucial role.

The second moment requires us to grasp reality as ever-changing and moving. Even though the human mind according to Bergson is used to regard everything in the manner of science as something static and the study of change without stopping the change and studying the change as change is uncharacteristic of it, even against the human intellect, it is possible. Our temporal nature and immediate experience of the duration of time make it possible. Duration is an essential part of intuition, because only inner experience of it, makes it possible to extend intuition to other objects of our knowledge. Intuition requires us to regard the object of our knowledge as a duration that also exists within us. With duration, we can

be "*in sympathy* with reality" (cf. Bergson 2007, 133, citation 20). When our inner duration is understood as analogous to the duration of the object of our study, we can grasp its nature and complete the understanding we already sketched out with scientific knowledge. Seeing things *sub specie durationis* means that we are regarding them through our own duration.

Philosophy is not solely the study of changing reality. For Bergson, philosophy itself is immersed in that change. Even the results of philosophy are temporal, they are not absolute and need to be improved by later philosophers. Simo Knuuttila, in his Keynote presentation at The Philosophical Society of Finland's Annual Colloquium 2022 in Oulu, posed the question of whether a philosopher's perspective on time can be anything but detached, limited to the "conceptual present" (*käsitteellinen nykyhetki*). Henri Bergson would answer, yes it can be, and it should be. According to Bergson, philosophy should abandon the perspective of eternity, unchanging "conceptual present". It should be temporal, finite, and speculative. For him, philosophy itself is temporal thinking, thinking in duration.

University of Jyväskylä

References

Bergson, Henri (1911). *Creative Evolution*. Translated by Arthur Mitchell. London: Macmillan & Co.

Bergson, Henri (1920). *Mind-Energy*. Translated by Herbert Wildon Carr. New York: Henry Holt and Company.

Bergson, Henri (1965). *Duration and Simultaneity*: *With Reference to Einstein's Theory*. Translated by Leon Jacobson. Indianapolis: The Bobbs-Merill Company, Inc.

Bergson, Henri (1972). *Mélanges: l'idée de lieu chez Aristote, Durée et simultanéité, Correspondance, Pièces diverses, Documents*. Edited by André Robinet. Paris: Presses Universitaires de France.

Bergson, Henri (1992). *Cours. 2, Leçons d'esthétique à Clermont-Ferrand ; Leçons de morale, psychologie et métaphysique au lycée Henri-IV*. Paris: Presses Universitaires de France.

Bergson, Henri (2001). *Time and Free Will: An Essay on the Immediate Data of Consciousness*. Translated by F. L. Pogson. Mineola, N.Y: Dover Publications.

Bergson, Henri (2007). *The Creative Mind: An Introduction to Metaphysics*. Translated by Mabelle Andison. New York: Dover Publications, INC.

Deleuze, Gilles (1988). *Bergsonism*. Translated by Hugh Tomlinson and Barbara Habberjam. New York: Zone Books.

Lapoujade, David (2018). *Powers of Time: Versions of Bergson*. Translated by Andrew Goffey. Minneapolis: University of Minnesota Press.

Lawlor, Leonard (2010). "Intuition and Duration: An Introduction to Bergson's 'Introduction to Metaphysics'." In, *Bergson and Phenomenology*, edited by Michael R. Kelly, 25–41. London: Palgrave Macmillan.

Lipsanen, Katariina (2021). "Bergson ja intuitio filosofian metodina" ["Bergson and intuition as the method of philosophy]". *Tiede & edistys 2021* (3): 133–46.

Tellier, Dimitri (2008). "Telle est ma vie intérieure, telle est aussi la vie en général." In *Annales bergsoniennes*, IV, 423–32. Paris: Presses Universitaires de France.

Verdeau, Patricia (2007). "Sur la relation de Bergson à Spencer." In *Annales bergsoniennes*, III, Bergson et la science, 361–75. Paris: Presses Universitaires de France.

Vieillard-Baron, Jean-Louis (2004). "L'intuition de La Durée, Expérience Intérieure et Fécondité Doctrinale." In *Bergson: La Durée et La Nature*, edited by Jean-Louis Vieillard-Baron, 45–75. Paris: Presses Universitaires de France.

Weber, Michel (2005). *La Dialectique de l'intuition Chez Alfred North Whitehead: Sensation Pure, Pancréativité et Contiguïsme (Introduction à La Lecture de Process and Reality (1929))*. Frankfurt: De Gruyter.

Weinstein, David (2019). "Herbert Spencer." *The Stanford Encyclopedia of Philosophy*. Fall Edition 2019.
https://plato.stanford.edu/archives/fall2019/entries/spencer/

Are We Triggering Injustice? Towards a Justice-Based Solution to the Non-Identity Problem

JOHAN SANDELIN

1. Introduction: Luck and Injustice

Anna lives a short, barely worthwhile life due to a genetic disease. Beatrice does so as well. But the cause of her short and barely worthwhile life is her foolishly exposing herself to risk for no good reason. Cedric lives a long, happy life through sheer good luck (example by Lippert-Rasmussen 2020, 9).

Some might think that the world would have been more just if Anna, Beatrice, and Cedric had lived equally good lives. Others might think that there is no injustice regarding these fates. But a common view in philosophy is that the world would have been more just if Anna and Cedric – but not necessarily Beatrice – had lived equally good lives. This being so because Anna had a difficult life through no fault of her own (see Lippert-Rasmussen, Spring 2023). In other words, we could say that Anna was born to an undeservedly bad faith, which was unjust or at least not morally good.

In relation to these kinds of fates, there are several important ethical questions. First, what responsibilities and obligations do individuals, societies and possibly even the world community have, to try to mitigate this sort of undeservedly bad fates? Second, what if anything, is it exactly that is unjust regarding these fates? Third, can it be morally acceptable to make decisions or act in ways that trigger more of these kinds of potentially unjust outcomes?

While the first of these questions – regarding the possible obligations to mitigate this kind of potential injustice – is important and interesting, it is unlikely that any solution to it

will be found, or any general agreement reached, in the foreseeable future. Luckily, this paper does not need to take any stance on this question.

The second question, regarding what the root cause is of Anna's fate being unjust (assuming that the intuition and common stance that Anna's fate is unjust is correct) is important for this paper. The example above suggests that it has to do with inequality or deficiency compared to other people or according to some yardstick – like what could reasonably have been expected.

The third question mentioned above was: Do we have a moral obligation *not* to act in a way that triggers these kinds of injustices? This question is central to this paper. It stretches towards the future, because it is in the future that the amount of unfair life fates is not yet determined and partly depends on our choices and actions.

As I already wrote, it is difficult to establish that we or our societies would have an ethical obligation to mitigate the (here assumed) unjust conditions many people are born into. Establishing that we have an ethical responsibility not to deliberately act in a way that could lead to this injustice increasing, might however be much easier and less problematic. This being so, because it is often assumed that we have a greater moral responsibility not to perpetrate, or be complicit in, injustices, than to correct those that already exist, and we have no complicity in. In other words, even if we could not establish that the societies have a moral obligation to correct or mitigate unjust fates such as those of Anna, it could become clear that it is morally wrong to deliberately trigger these kinds of outcomes and thus become complicit in them.

When we consider this question, we are simultaneously faced with another important question: Is this possible injustice independent of whether those who are subjected to it are harmed or not? Can we say that it was unjust that Anna was born into a difficult condition, even if this did not harm her in the sense that she, because of it, was worse off than she otherwise would have been? And if so, can we go on, arguing that someone acted morally wrong if he or she acted in a way that increased the risk that Anna, or those who could have been born in her place, would be born to an unfair fate? This

question is important because the answer to it can have significant implications in ethics.

In ethics, there are some relatively broadly accepted basic moral principles that are thought to determine whether our actions are morally wrong. According to one of these principles, an act that does not harm anyone does not wrong anyone. According to another of these principles, an act that does not wrong anyone is not morally wrong. And according to one more such principle, an act harms someone only if this person is worse off than he or she otherwise would have been (for these principles, see Boonin 2014, 3-5).

While these principles appear to generate generally plausible results regarding issues affecting currently living people, they have been shown to produce counterintuitive results regarding our responsibilities to future generations. For example, we have not with the help of these principles, been able to satisfactorily explain why it is morally wrong of our generation to live in a way that leads to global warming and difficult conditions in the future, as this lifestyle probably also affects the identity of future people so that no one born into a hot difficult world could have been born into better conditions. This is an ethical conundrum, called *the Non-Identity Problem that* has remained unsolved for half a century (see Parfit 1983 for its introduction).

If we can show that Anna's fate was unjust even though she could not have got a better life – because this injustice is based on a deficiency according to some relevant comparative standard that is independent of any counterfactual assessments of how good her life could have been – we can perhaps also show that it is morally wrong to act in a way that deliberately triggers this injustice. Then we might have a solution to the Non-Identity Problem at sight.

I will thus try to answer the question of whether a focus on naturally occurring injustice, and how we affect the amount of this injustice, could help us solve the Non-Identity Problem. The idea is that an action can be morally wrong on two different possible grounds. An action can be morally wrong, as previously thought, because it harms someone according to the counterfactual account of harm, but an action can also be morally wrong because it deliberately causes a person to be born into an unjustly bad fate. Both harm and justice are

thus thought to matter morally. I will call a developed form of this stance the justice-based theory or principle.

According to David Boonin (2014, 19-28), a successful solution to the Non-Identity Problem must show that the Implausible Conclusion – that we are not doing moral wrong in identity affecting choices like causing climate change – can be avoided. A successful solution should also meet his independence-, modesty- and robustness requirements. I.e. the solution should not be based solely on the fact that it would enable us to avoid the Implausible Conclusion (the Independent Requirement); the solution should be sufficiently robust to be able to avoid the Implausible Conclusion even if one of the premises leading to the Implausible Conclusion is modified (the Robustness Requirement); and the solution should not have implications that are even more implausible than the Implausible Conclusion (the Modesty Requirement). In the third section, I will briefly discuss how the here suggested Justice-Based Principle fares regarding some of these requirements.

This paper should be seen as a first tentative framing of a question and theory, rather than a comprehensive final solution to the Non-Identity Problem. But before I go into this theory, I will start by presenting the Non-identity problem in more detail as well as briefly discuss three types of established approaches for solving the Non-Identity Problem and why they all seem to fail.

2. Traditional attempts for solving the Non-Identity Problem

The Non-Identity Problem comes in two versions, so-called direct and indirect versions. The most typical examples of the indirect version of the Non-Identity Problem concern climate change and the pursuit of a general policy of Depletion. In his last attempt to solve the Non-Identity Problem – almost forty years after he first discovered it – Derek Parfit (2017, 122-123) gave such an example:

> Suppose […] that we and the other members of some large community could choose between two energy policies, one of which would be cheaper but would increase global warming, thereby having various effects that would greatly lower quality

> of life that would be had by very many people in several later centuries. Some of the effects of our policy – such as floods, draughts, heat waves and hurricanes – would kill many of these future people. Despite having these effects, our choice of this energy policy would not be worse for any of these people, not even those who would be killed. If we had chosen the more expensive policy, which would not have had these bad effects, these future people would never have existed, and their non-existence would not have been better for them. It would have been different people who would have existed instead and lived much better lives.

Parfit is here assuming the nowadays generally acknowledged circumstance, that policies which have profound impacts on people lives – where they work, live, whom they meet and ultimately when and with whom they have children – affect the identities of the future people. A process that with Parfit's words will not dissolve in time, like rings in the water, but instead is self-strengthening (Parfit 2017). Since the choice of the cheaper policy, in the example above, would be worse for no one, the challenge, of the indirect version of the Non-Identity Problem, is to explain why it would be morally wrong to choose it.

An example of the direct version of the Non-Identity problem is the following by David Boonin (2014, 2):

> Wilma has decided to have a baby. She goes to her doctor for a checkup and the doctor tells her that there is some good news and some bad news. The bad news is that as things now stand, if Wilma conceives, her child will have a disability. The doctor cannot say precisely what the disability will be, but he can tell that while the disability will be considerably far from trivial, the child's life will nonetheless clearly be worth living. The disability will be irreversible. There will be no way to eliminate it or to mitigate its effects. The good news is that Wilma can prevent this from happening. If she takes a tiny pill once a day for two months before conceiving, her child will be perfectly healthy. Fully understanding all of the facts about the situation, Wilma decides that having to take a pill once a day for two months before conceiving is a bit too inconvenient and so chooses to throw the pills away and conceive at once. As a result of this choice,

> her child Pebbles is born with a significant and irreversible disability.

The challenge in this version is to explain why Wilma's choice would be morally wrong even though Pebbles owes her existence to Wilma's choice; she could not have got a better life and is glad to be alive.

Several different solutions to the Non-Identity Problem have been proposed during the last decades, but none of these have been considered satisfactory.[1] In this section, I briefly discuss three types of established approaches for solving the Non-Identity Problem that I call the "Idealistic Approach", the "Harm-Based Approach" and the "Capitulation Approach".[2] I then go on to argue that all these approaches fail for several different, generally acknowledged, reasons.

2.1. The Idealistic Approach

Derek Parfit (1984, 367) had a clear view on the question how the Non-Identity cases affect our moral obligations:

> We may be able to remember a time when we were concerned about effects on future generations, but had overlooked the Non-Identity Problem. We may have thought that a policy like Depletion would be against the interests of future people. When we saw that this was false, did we become less concerned about effects on future generations? When I saw the problem, I did not become less concerned. And the same is true of many other people. I shall say that we accept the *No-Difference View*.

---

[1] For readers who are interested and not familiar with The Non-Identity Problem, its attempted solutions and the problems with these, David Boonin's (2014) book *The non-identity problem and the ethics of future people* is recommended.

[2] The name "Idealistic Approach", because it is here assumed that benefitting is morally (almost) as important as not harming. The "Harm-Based Approach" because these attempts try to solve the Non-Identity Problem by some revision of our understanding of harm. The Capitulation approach, because this "solution" to the Non-Identity Problem stems from Boonin's decision to give up the hope of finding a way to solve the Implausible Conclusion that Wilma does not do anything wrong by not taking the Pill and conceiving disabled Pebbles (see the citation in section 2.3)

Parfit (1984, 369) thought that our obligations towards the future are just as strong despite the fact of Non-Identity. Initially he thought that, in order to solve the Non-Identity Problem, we should appeal to an impersonal ethical theory and began by considering a principle Q:

> Q: If in either of two possible outcomes the same number of people would ever live, it would be worse if those who live are worse off, or have a lower quality of life, than those who would have lived.

Since Q only concerned "same number choices" and since many of our choices and actions will not only affect the identities of future people but also their number, the question arose: How important is the quality of life of each future person compared to the number of future people that will exist? For example, would it be better if a thousand people with a very high quality of life were born on a remote island in the future, or, if instead, ten thousand people with somewhat lower quality of life were born on the same island in the future, everything else being equal?

Parfit (1984, 386-387) suggested two basic principles that could answer this question and possibly replace Q:

> *The Impersonal Average Principle*: If other things are equal, the best outcome is the one in which people's lives go, on average best.
>
> *The Impersonal Total Principle*: If other things are equal, the best outcome is the one in which there would be the greatest quantity of whatever makes life worth living.

Both principles aspire to solve the Non-Identity Problem by showing why suboptimal choices regarding the future are morally wrong, even if these choices affect the identities of future people. However, both principles also imply different very counterintuitive conclusions. The Impersonal Average Principle implies what I call the "Dire Conclusion" according to which creating people with negative lives can be morally right in extremely bad situations – where those already living have even worse lives, so that this would increase the average quality of life. The Average Principle also amounts in what I call the "Discriminatory Conclusion": As people being born

in countries with relatively low quality of life likely will decrease the worldwide average quality of life, it would be morally better if the people in these countries would abstain from having children. Actually, as Parfit wrote (1984), according to the Average Principle, it would, counterintuitively, be best if only one very happy person existed.

The Impersonal Total Principle has, on its part, been found to imply the "Repugnant Conclusion" according to which creating an enormous future population with life only barely worth living would be preferable – if the total amount of quality of life in this outcome would be the highest due to the population's incredible size (Arrhenius, Ryberg & Tännsjö 2022). Furthermore, impersonal ethical principles have become increasingly abandoned since the 1980s, as utilitarianism has become less popular and has been criticized exactly for being an impersonal theory which treats persons as mere containers of utility that are replaceable and do not in their own right matter.

In his last attempt to solve the Non-Identity Problem, Parfit (2017, 123-124) had revised his view and wrote:

> When I first defended the No Difference View, I made what I now believe to be a bad mistake. I suggested that, when we consider the cases that raise the Non-Identity Problem, we should appeal to principles that are *impersonal* in the sense that they do not appeal to facts about what would affect particular people for better or worse.

He then suggested a person-affecting view that would maintain the comparison aspect between the alternative future people we can bring about. Even if it would be true that we cannot harm future people by bringing them into flawed existences, he argued that our actions can be morally wrong by not benefitting people as much as we easily could. He wrote that following Jeff McMahan (2013, 6-7) we can claim:

> If someone is caused to exist and to have a life that is worth living, that is good for this person, giving him or her an existential benefit.

Parfit (2017, 137) then continued:

> To avoid giving people reasons to complain about what we have done to them, it is enough to do nothing that would be bad for these people. We could achieve this moral aim in a purely negative way, by doing nothing. Though non-maleficence is negative, beneficence is positive. We can give people reasons to be grateful only by doing things that are good for these people. And though it is important not to act badly it is also morally important to act well.

This led him to formulate the following:

> *The Wide Principle:* One of two outcomes would be in one way worse if this outcome would be less good for people, by benefiting people less than the other outcome would have benefited people.

However, the Wide Principle again led to the question of how to aggregate quality of life. What matters morally? Is it what in the collective sense together benefits people more, or is it what in the individual sense benefits each person more? The collective perception would, as Parfit noted, restate the Total Principle in person-affecting terms. The individual perception would restate the Average Principle in person affecting terms. Since Parfit was aware of the problems with both views, he suggested a combination of them, which he hoped would be more balanced:

> *The Wide Dual Principle:* One of two outcomes would be in one way better if this outcome would together benefit people more, and in another way better if this outcome would benefit each person more.

Unfortunately, Michal Masny (2020) has shown that even though more balanced than his earlier attempts, Parfit's Wide Dual Principle still ends up in the Repugnant Conclusion. Since the Repugnant Conclusion is at least as implausible as the Implausible Conclusion, the Wide Dual Principle does not qualify as a satisfactory and unproblematic solution to the Non-Identity Problem.

2.2. The Harm-Based Approach

Not everyone agrees with Parfit, and some have chosen a different strategy to solve the Non-Identity Problem, commonly

known as the Harm-Based Approach. This approach avoids the assumption that we have an almost as strong moral obligation to benefit people as we have, to not harm them. But this advantage comes at a cost: The Harm-Based Approach must only base the moral judgment of our choices on the people born to suboptimal lives and not on any comparison with other people. It must show that these people have been harmed even though they have received the gift of life and are not worse off than they otherwise would have been, have lives that are clearly worth living and are glad to be alive.

Robert Huseby (2012) defends a version of Telic Sufficientarianism. According to this version, a person who is born into a life that is below a certain sufficient threshold of welfare is harmed even if this person is above the neutral level where life ceases to be worth living. Shortfalls from a sufficient level of welfare are morally bad, and this is all that matters welfare-wise. In addition, it is in itself bad if a person is not sufficiently well-off and worse the further from a sufficient level a person is and worse the more people are not sufficiently well off (Huseby 2012).

Molly Gardner (2019) has proposed an existence account of harming:

> *Harm* (def.): A state of affairs, T is a harm for an individual, S, if and only if
>
> There is an essential component of T that is a condition with respect to which S can be intrinsically better or worse off; and
>
> If S existed and T had not obtained, then S would be better off with respect to that condition.

As Gardner (2019, 433-438) explains, "the account asks us to compare the world containing both the victim and the alleged harm, to a world where the victim exists without the alleged harm; we then check whether, in the latter world, the victim is better off". If the answer is yes, the event harmed the victim.

Unfortunately, both Huseby's and Gardner's versions of the Harm-Based Approach end up in very counterintuitive conclusions. One such is the "Fatal Conclusion" that there would be nothing wrong with causing humanity to die out. As Parfit (2017, 118) did, we can imagine a situation where

we discover how we can live for a thousand years, but in a way that makes it impossible for us to have children. According to the Fatal Conclusion, there would be nothing wrong with choosing this option.

Huseby's Telic Sufficientarianism fails to explain what is wrong with the Fatal Conclusion, because no one in an empty world would be below a sufficient level of welfare. In addition, according to Telic Sufficientarianism, it may even be the best choice to let humanity die out – if we assume that otherwise there will always be some people who are below the sufficient level (see also Boonin 2014, for this critics). The Existent Account of Harming also fails this test, as there would be no victims with any alleged harms in a world without humans.

Another counterintuitive conclusion that these harm-based approaches fail to avoid, is the already mentioned "Discriminatory Conclusion". According to this conclusion, the citizens of countries with low quality of life would do something wrong by having children. Telic Sufficientarianism seems unable to avoid this conclusion without ending up in another unintuitive conclusion.

To see why this is so, consider the following. In some countries the quality of life as well as life expectancy is low and will very likely remain low in the foreseeable future. In fact, we can plausibly assume, that the average quality of life in these countries will not in the foreseeable future be higher than that of Pebbles quality of life in the example above – at least if we assume that Pebbles disability is not particularly severe. We then face the following dilemma: If we want to show that Wilma did something wrong by not taking the pill and conceiving disabled Pebbles, we must raise the sufficientarian threshold so much that Pebbles is below it. But then also those living in the poor countries will do something wrong by having children – and we end up in the discriminatory conclusion. To avoid this, we must lower the threshold – but alas, then we fail to show that Wilma's act was morally wrong and instead end up in the "implausible conclusion" according to which almost anything goes regarding the future.

Defenders of the harm-based approach might argue that this approach is not vulnerable to the discriminatory conclu-

sion, since the bad circumstances that children born into poor countries will face are usually *not caused* by their parents. However, I do not believe that this defense is tenable. It is quite clear that, for example, it is about as morally wrong to give one's child a bad upbringing as it is to let someone else take care of one's child if one is sure that this person will give the child an equally bad upbringing – this even though in the latter case, one would not directly "cause" this bad upbringing. In the same way, the parents in the poor countries know the circumstances into which their children will be born, and they are the reason for and the cause of their children's birth.

Sufficientarian theories have also been criticized on the grounds that the threshold, or level, is arbitrary and too sharp: it is not intuitive that causing someone to be slightly above it would not be wrong at all but making someone to be slightly below it would be morally wrong. Boonin (2014) has further given several examples where the Sufficientarian theories seem to generate the wrong answers and the counterfactual theory of harm the right ones. In response to this, revised theories with more than one threshold have been proposed, as well as some versions that include both a sufficient level and the counterfactual account of harm. However, such ad hoc hybrid theories are not intuitive and do not escape all of the problems mentioned above. They might cause new kinds of problems and I do not find it likely that they could offer a satisfactory solution to the Non-Identity Problem.

The Existence Account of Harming also seems to fail to avoid the discriminatory conclusion. To conceive children in countries with difficult living conditions would be morally wrong, according to the theory, since these children would be worse off than they would have been in an ideal world where nothing would have weighed down their quality of life (see Boonin 2019 for further criticisms).

2.3. The Capitulation Approach

In 2014, David Boonin wrote in his book *The Non-Identity Problem and the Ethics of Future People*:

> I have been thinking of and about the non-identity problem for a number of years now. When I first encountered the problem, I assumed that there must be a satisfactory way to avoid the Im-

> plausible Conclusion by modifying one of the premises of the argument that gives rise to it. The only question seemed to be which premise to target and how, precisely to modify it. After a number of years of false starts and dashed hopes, though, it occurred to me one day to wonder whether the most reasonable response to the problem might simply be to give up and bite the bullet.

I call this approach for solving the Non-Identity Problem, the "Capitulation Approach" because it does not try to explain why Wilma's choice would have been wrong. Instead, this approach assumes that choices that lead to a lower quality of life in the future are not morally wrong if these choices affect the identities of future people, so that no future person has a worse life than he or she could have had. Boonin calls this counterintuitive assumption the Implausible conclusion and is ready to accept it. If Boonin were right, we would not have to find a new ethical principle that would replace the counterfactual account of harm. The Non-Identity Problem would dissolve because there would simply be no other problem than that our intuitions are wrong regarding our responsibilities to the future.

Other ethicists have found it difficult to accept that a theory could be right if it ends up in the Implausible Conclusion – which implies that when it comes to future people almost anything goes – as most actions that profoundly affects quality of life in the future also likely will affects the identities of the future people. Furthermore, the Capitulation Approach also fails the test to avoid the "Fatal Conclusion" according to which there would be nothing wrong by causing humanity to die out. This being so since no one can be worse off than he or she would otherwise have been by not being born.

		Counterintuitive conclusions that should be avoided:				
		The Implausible Conclusion: When it comes to future people almost anything goes	**The Fatal Conclusion:** There would be nothing wrong by causing humanity to die out	**The Dire Conclusion:** Creating people with hellish suffering can be a moral duty	**The Repugnant Conclusion:** An enormous population with a life that is barely worth living is preferable	**The Discriminatory Conclusion:** People in countries with relatively low quality of life should not have children
Capitulation Approach	**The strong Narrow Deontic Principle**	FAIL	FAIL	PASS	PASS	PASS
Idealistic Approach	**The Average Principle**	PASS	PASS	FAIL	PASS	FAIL
	The Total Principle	PASS	PASS	PASS	FAIL	PASS
	The Wide Dual Principle	PASS	PASS	PASS	FAIL	PASS
Harm-Based Approach	**Telic Sufficien-tarianism**	PASS	FAIL	PASS	PASS	FAIL
	The Existence Account of Harming	PASS	FAIL	PASS	PASS	FAIL

Table 1. *Challenges for different established principles developed to solve the Non-Identity Problem.* The table illustrates how the principles manage or fail to avoid some very counterintuitive conclusions. FAIL, as opposed to PASS, means that the principle is incapable of explaining why the specific counterintuitive conclusion is wrong.

3. A justice-based solution to the Non-Identity Problem

Humans (and animals) are born into vastly different conditions due to partially random processes that they cannot control themselves and cannot be held accountable for. Some are born with intact senses and good health into an environment where they have the conditions to live full, good lives, while others are born with sensory issues, poor health or into an environment where they have negligible opportunities to live other than limited and tolerable lives which still, however, can often be considered worthwhile. Since those born with relatively bad fates are born with these fates through no fault of their own, and there is nothing that can morally justify

people's different good initial life conditions, these differences can be considered unjust, or at least undeservedly bad. This is the stance of this paper.

The extent, to which societies have a responsibility to try to neutralize this inequality, and compensate victims of undeservedly bad fates, has long been debated in ethics. However, reaching an unequivocal answer to this question does not seem possible, at least in the foreseeable future. A related question that has received less attention is what our responsibility is for actions that trigger and indirectly increase these kinds of comparatively unjustly bad fates. Are such actions unjust and morally wrong? Establishing that we have an ethical responsibility *not* to act in a way that leads to an increase in the proportion of victims with comparatively unjustly bad fates may be easier and less problematic than establishing that we would have a responsibility to compensate victims whose bad fates we are not guilty of. This being so, because it is often assumed that we have a greater moral responsibility not to perpetrate, or be complicit in, injustices, than to correct those that already exist, and we have no complicity in.

In this article, I assume that it is morally wrong to trigger these kinds of unjustly bad fates and that if you do so, you become complicit in this injustice. The question I ask is whether such a theory – or moral principle – could solve the Non-Identity Problem.

A tentative version of this theory and principle can be formulated as follows:

> JP: There are good reasons to hold that an action X is morally wrong if either:
>
> – X makes a person worse off than he or she otherwise would have been (the conventional counterfactual account of harm); or
>
> – X deliberately causes a person to end up in a comparatively unjustly bad condition, and this could reasonably easily have been avoided and was not necessary to avoid the person's condition becoming very dire.

This justice-based principle (JP) can be favorably compared to harm-based theories that assume some sufficientarian level and which I discussed in the previous section. Both JP and

the sufficientarian theories assume that Wilma's action is morally wrong because her action results in Pebbles being born into a state that is somehow too bad or problematic. According to the sufficientarian theories, Pebble's condition is too bad in welfare terms – it is below a certain level, and according to the principle of justice proposed above (JP), Pebble's condition is unjust because it is comparatively too bad and there is nothing that justifies this.

I believe that the Justice-Based Principle has several advantages over the sufficientarian theories. It avoids the problems of having a theory with several different, possibly conflicting understandings of what harm entails. It is not based on a non-comparative sharp level that is arbitrary, but instead is based on a comparison with other people or with what could be reasonably expected.

At the same time, I am grateful to a reviewer of this paper for pointing out some potential weaknesses with the Justice-Based Principle. One such is how the Justice-Based Principle manages to escape the above-presented Fatal-, Discriminative- and Dire Conclusions. Indeed, it seems difficult to explain how JP could avoid the Fatal Conclusion, since no future humans could have been wronged in a world without any future humans. It seems that this empty world challenge poses a powerful conundrum for all person-affecting principles – except for Parfit's Wide Dual Principle which takes into account the moral significance of benefitting.

However, it is possible that it is easier to explain why it would be morally wrong to allow humanity to die out using fairness and justice-based arguments than using harm-based arguments. Since previous generations have gone to great lengths to ensure the prosperity of future generations, intuitively, it seems that it would be unfair towards the previous generations to allow humanity to die out (see for example Thompson 2016, 289-300). It is not as easy, in a similar way, to explain how we would harm the previous generations (in welfare terms) by letting humanity die out, as the previous generations do not exist anymore and therefore are not easily harmed. Because of this, I believe that JP, regarding the Fatal Conclusion, could have an advantage over the harm-based arguments. But to avoid JP ending up in the Fatal Conclusion, with arguments based on fairness or justice towards the pre-

vious generations, would require a development of JP which I cannot present here.[3]

Regarding the Discriminatory Conclusion, I believe that the Justice-Based Principle can escape it because people born in countries with a low quality of life do not get a life that is unfair compared to their fellow human beings in these countries or to what could reasonably have been expected. If this is right, then JP is not discriminating against these communities and people. In chapter 3.2, however, I discuss in more detail the question of how to determine what we should compare with when we evaluate whether an individual's fate is unfair or not.

Regarding the Dire Conclusion, it could be suspected that the Justice-Based Principle would be unable to escape this counterintuitive conclusion, because of the following reason: If the relevant comparative society has a very low – or even negative – level of welfare, causing a person to end up in a comparatively only slightly better condition, than that of the society, could *not* be unjust.

While it is true, that JP might have difficulties in recognizing anything wrong with this act on grounds of justice, it is possible that the other part of the JP account – the counterfactual harm-criteria – can explain why it would be wrong to cause someone to be worse off than nothing, or to have negative welfare. This being so, as a person that is born into a life worse than nothing could be understood to be worse off than he otherwise would have been.[4] Neither of JP's criteria for a morally wrong action is met, however, if a person's situation is not made worse than it otherwise would have been, and the person's situation is barely positive but not worse than that of his fellow humans or what could reasonably have been expected. One such imaginable case is if the citizens of a

---

[3] Another option, for dealing with the Fatal Conclusion, would be to show that something else matters morally and can be morally wrong than harming and being unfair – which would save both the harm-based theories and JP from the Fatal Conclusion, but probably would imply new moral dimensions and new kinds of problems related to these.

[4] Recently, Olle Risberg (2023) has, however, in an interesting paper argued that the counterfactual account of harm is unable to explain why causing someone to have a life worse than nothing would be morally wrong if we assume actualism.

society with very difficult, but positive, living conditions choose to have children who are likely to have difficult but worthwhile lives. Here I, however, think that JP generates the right result and avoids ending up in the Discriminative Conclusion. This is because it would be difficult to explain why parents in such countries would do something morally wrong by having children.

One more challenge with JP is according to the one reviewer, that it could be too broad and make wrong every action that aim to improve the general condition of people in a society (which can already have a high quality of life) but do it in an unequal way so that e.g. a single person's condition will not be improved and he will remain at the original, very high but significantly lower level than others, so that he will end up in a comparatively unjustly bad condition as the result of the improvement of others.

This is an interesting comment and challenge. I believe that a policy which deliberately means an improvement for all but a few people should strive to compensate these few unfortunate people if possible, so as not to be unfair to them and on these grounds be morally ambiguous. If it is not possible to compensate them, I believe, in line with JP, that it could be morally better not to carry through with the policy or even morally wrong to do so. As an analogical comparison we can think about a party where all the guests get very good food. They also get a good dessert, but now you are offered the opportunity to treat all but two of the guests to a truly excellent cake for free. I think most of us would feel it would be morally wrong to accept this offer as the situation is already very good for all the guests and this would be unfair to two of them. This position is in line with the Justice-Based Principle. On the other hand, actions or policies that strive for general improvements and are not associated with increased inequality, I don't think can be judged as unfair if certain individuals as an unfortunate consequence (bad luck) get a comparatively worse situation compared to their fellow human beings than before. In this case, it is an unfortunate consequence that could not be predicted and which, according to JP, does not mean that the action or policy would be morally wrong. But even in these cases, compensation, if possible, could be mor-

ally desirable. Thus, all in all, I do not think that this criticism could show that JP would be too broad.

One thing, in addition to those mentioned earlier, that instead *supports* the justice principle (JP) is that the counterfactual account of harm has increasingly come under scrutiny and seems to be unable to account for all kinds of wrongful actions. One such criticism is that some actions appear to be unjust, and on these grounds morally wrong, even though they do not make the person they affect worse off than she otherwise would have been. Such actions are e.g., exploitation, humiliation, dishonesty, and lack of solidarity (Piazza et al. 2018). A common example of unfair exploitation is for example the following:

> *Thirsty in the desert*: A man meets a woman in the desert who is dying of thirst. The man has several bottles of water. He either offers to sell a bottle of water at an unreasonably high price or wants in exchange a kissing session with the woman (See for example Zowlinski 2022).

In both cases, we can assume that the woman accepts and that she is not worse off than she otherwise would have been and happy that the man came her way. But at the same time, the man's action seems clearly morally wrong because he took advantage of the woman's vulnerable position and treated her unfairly – or caused her to end up in an unjust condition, which according to the here proposed principle (JP) is morally wrong. Another example is

> *The sadist and the beggar*: a hungry beggar sit on the street and a sadistic man simultaneously gives the beggar money and humiliates him.

Once again, it might be that the beggar is not worse off welfare-wise than he otherwise would have been all things considered, yet the man acted unfairly, and on this ground, morally wrong.

All of this seems to support the assumption that not only harm should be considered regarding what actions are morally right/wrong but also injustice. This supports the proposed justice-based principle (JP). But as the reader can imagine, there are several challenges and difficulties related to the principle not yet discussed.

3.1. Boonin's criticism of fairness-based solutions to the Non-Identity Problem

In his book *The non-identity problem and the ethics of future people*, Boonin considers some fairness-based solutions to the Non-Identity Problem. He focuses mainly on attempts to use Rawls's veil of ignorance to develop such a solution. He however shows that these attempts fail for several reasons and writes (2014, 130):

> If Wilma conceives Pebbles rather than Rocks, to whom can her act be unfair? Her act isn't unfair to herself or to Pebbles or to any other actual person, because deliberators representing the interests of all of those people would endorse a rule permitting Wilma to conceive Pebbles. And how can Wilma's act be unfair to Rocks? Rocks does not exist, and an act can't treat a person unfairly if that person never exists.

Boonin also considers attempts to solve the non-identity problem by showing that those born into suboptimal lives have been unfairly exploited. He rejects this attempt on the grounds that a criterion for, and central aspect of, exploitation is that the exploiter could easily have offered the exploited a better deal, whereas Wilma could not have offered any other deals to Pebbles (Boonin 2014, 142).

I agree that The Non-Identity Problem cannot likely be solved with a theory of exploitation because the actions that lower the quality of life in the future and at the same time affect the identity of future people do not seem to fulfill the characteristics of what is generally considered to be exploitation. At the same time, however, I believe that an action by a person X can be unfair to a person Y even if X could not have offered any other deals to Y – except non-action – and Y is not worse off, because of the action, than Y otherwise would have been.

To see why this is so, reconsider the example of exploitation above and the *sadist and the beggar*. We can hypothetically imagine that the men in the examples cannot act in any other ways than to humiliate the beggar or offer an unfair deal to the woman in the dessert, if these men act at all. We can also assume the woman in the desert and the beggar are not worse

off if the men decide to act. But despite this, these actions do not seem morally justified. This is because I think that *benefiting does not justify injustice*. *We cannot reasonably hold that we have the moral right to treat someone unfairly just because we benefit this person at the same time*, even under these kinds of conditions. The men's actions are morally wrong, at least if the men have not before their actions took place, explained the situation to the woman and the beggar – their inability to help without acting unjustly – and the woman and the beggar have consented, fully understanding the situation. The only other thing that I think could justify the actions of the men is if the woman and the beggar would otherwise almost certainly have been close to dying[5].

That one cannot act better towards a person does not therefore mean that one's action would surely be morally acceptable, even if the person in question, all things considered, would be better off following the action. The action can still be morally wrong because it is unjust. We therefore cannot rule out the possibility that the Non-Identity Problem could be solved with a theory of justice.

3.2. Further challenges with the justice-based principle (JP): What should we compare with?

An obvious problem with the here proposed theory and justice principle (JP) is what we should compare with. The principle states that an action by a person X can be morally wrong if X causes a person Y to end up in a comparatively unjustly bad condition. But how are we to determine whether person Y's situation is comparatively unjustly bad?

To answer this, I will first examine Boonin's example with Wilma and Pebbles. In Figure 1, I assume that Wilma belongs to a society where the quality of life is medium. That the quality of life is, and throughout history has been, different in different societies is a fact that several studies on quality of life, happiness, and subjective well-being have shown. There

[5] Saving someone from dying or nearly dying is a very different thing than not bringing a person into existence. This is because it cannot harm or be bad for a person not to come into existence when this person does not exist.

are certain societies where the quality of life is higher and other societies where it is lower.

Let us suppose that Pebbles belongs and is born into a medium quality of life society but that her life, because of her disability, is significantly lower than the average quality of life in her society. How should we decide whether Pebble's fate is comparatively unjust? Shall we compare Pebbles' quality of life with Rock's quality of life, Wilma's quality of life, Pebble's society's quality of life, or with the quality of life in societies with a higher or with a lower quality of life than Wilma's society, or with the quality of life in her society a few generations back in time? These comparisons give different results as we see in Figure 1.

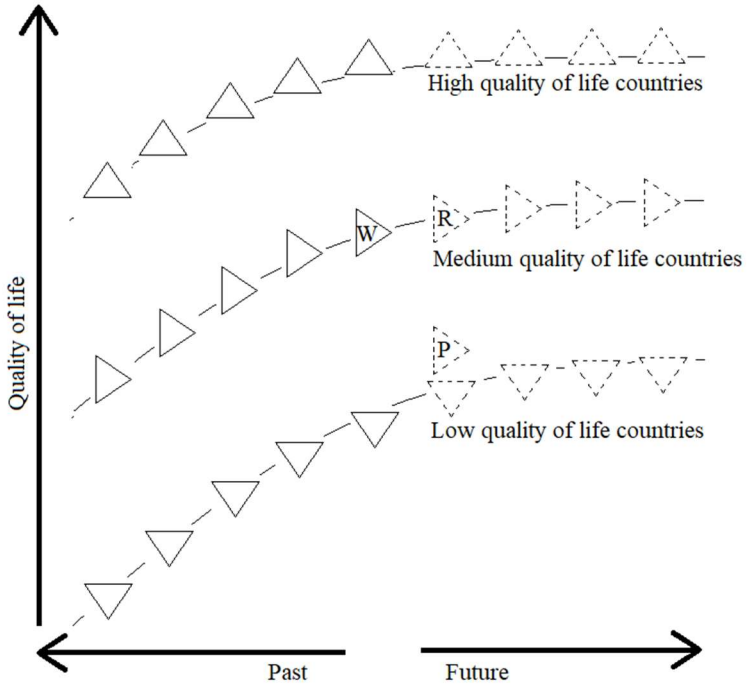

Figure 1. *Whom or what should Pebbles' fate be compared to?* In this figure, Pebbles is supposed to be born in and belong to a society

with medium quality of life which is characterized by the symbol ▹. W stands for Wilma. Low quality of life countries is represented by the symbol ▿ and high quality of life countries by the symbol ▵. P stands for Pebbles and R stands for Rocks. Pebbles is, in the figure, below the line of medium quality of life countries even though she is born in one of these countries, because her quality of life is much lower than the average in medium countries – in fact her quality of life is near the average quality of life of those countries that have low quality of life.

If we Compare Pebbles' fate to Wilma's fate, Rocks' fate, or Pebble's own society, then Pebble's fate is comparatively unfairly bad. But if we compare Pebble's fate with previous generations or the quality of life in societies with a low quality of life, Wilma's choice has not led to Pebble's fate being comparatively unfairly bad (see figure 1). How should we then be able to determine what it is reasonable to compare with? Does this mean that JP is arbitrary and can't meet Boonin's requirements for among other things modesty? According to this requirement a solution to the Non-Identity Problem should not have implications that are even more implausible than the Implausible Conclusion (see the introduction section).

Moral considerations regarding to what extent individuals, who get comparatively bad fates through no faults of their own, should be helped and compensated, are usually limited to the citizens or residents of a given society (See Elkins 2007). Though this limitation has been criticized, I believe that the same limitation is reasonable when evaluating whether people's fates are comparatively unfairly bad. It is unfair that a person who belongs to a certain group or society receives much worse conditions than the others in this group or society, regardless of the conditions outside this group or society. It is certainly also true that it is unfair, or at least undeserved, that the citizens of some societies have a much lower quality of life than the citizens of other societies, but this unfairness is not something Wilma increases or decreases with her choices and not something she can be held accountable for. Therefore, this broader inequality should not be the primary consideration when establishing the possible wrongfulness of Wilma's action.

But even if we make this limitation, the question remains: Should we compare Pebbles' fate with Wilma's, Rocks' or with her society's or with all of them. All these comparisons seem to generate the same results, but that doesn't mean the question is irrelevant. Comparing Wilma's fate to Rocks' may seem implausible as Rocks never existed, but I think Rocks' as well as Wilma's and Pebbles' society can all serve as relevant counterpoints – *they all show what should be expected of Pebbles' quality of life*. And one possible interpretation is that it is precisely this we should compare with when we evaluate if Pebbles' fate is unfairly bad – with what would reasonably have been expected regarding the fate of Wilma's children.

To see if this is a reasonable assumption, I shall continue by considering the indirect version of the Non-Identity Problem and Parfit's example with global warming. In this example, the future people are born into a Hothouse Earth due to the current generation's inability to make sufficient sacrifices and implement policies that would reduce greenhouse gas emissions. However, the existence of these future people is dependent on this inability of the current generation and therefore we cannot show, with the help of the counterfactual account of harm, that the current generation would have harmed these future people. When we examine the example with the principle JP, the question is then whether the fates of the future people are comparatively unfairly bad, which would show that the behavior of the current generation was unjust and morally wrong. To determine whether the fates of the future people is comparatively unjustly bad, it is again crucial to know what their fate should be compared with.

Above, we saw that a comparison with closely related people or groups or citizens of the same society seemed more plausible than a comparison with less closely related people or citizens from other countries. From this point of view, one could assume that the most relevant comparison, also in this example, would be one between the future humans among themselves. However, such a comparison is clearly not plausible: If we were to compare the fates of future people with each other, then their fates would not be unfairly bad – however bad they would be – because all would share the same difficult lot. In the example with Wilma and Pebbles, the comparison with individuals from the same generation and

society was relevant because the question concerned whether the actions of a single individual, Wilma, were morally wrong, and the behaviors of the other individuals in her society were assumed to be morally right. In this example, however, the question is whether the whole society has acted morally wrongly towards the future people, and therefore the comparison must be broadened and apply to people who are not assumed to have been exposed to unfairly bad fates. Or to put it another way: comparing the fates of the future humans with each other would be as misleading and as irrelevant as comparing the fate of Pebble with the fate of her twin sister if we assume that Wilma would have had twins, and both would have been born with a disability.

What then is a reasonable yardstick to compare with? The next closest relevant comparison seems to be the nearest previous generation that is not assumed to be victims of unfairly bad fates – in other words the current generation. Here too we can assume that a reasonable comparison is limited to the same societies. The question that should be asked then is whether the fate of future people is unfairly bad compared to the fate of the current generation in the same societies, and that seems to be clear. If the current generation has a lifestyle that results in a Hothouse Earth for the future people, then the future people through no fault of their own have received comparatively unfairly bad living conditions. Furthermore, if it is the choice of the current generation that has resulted in this unfair relationship – as it would be – then this means that the current generation has acted unfairly and, on this ground, morally wrong.

But also in this case, we can imagine that the yardstick that we compare to – or should compare to – is what quality of life the current generation could reasonably be able to offer the people of the future without making too many sacrifices.

4. Conclusions and some considerations

The Non-Identity Problem constitutes one of our times biggest conundrums in philosophy. Before this problem is solved, we do not have a coherent ethical theory regarding our moral responsibilities towards the future generations,

and philosophy's ability to help humanity cope with perhaps its greatest challenge remains seriously compressed.

In this paper, I have tried to make a first tentative investigation into the question of whether The Non-Identity Problem could be solved with a justice-based theory that is partly based on our moral responsibility not to act in a way that leads to future people having comparatively unfairly difficult fates.

The proposed Justice-Based Principle seeks to solve The Non-Identity Problem by evaluating Wilma's action using two different yardsticks – one counterfactual and one justice-based: Firstly, a comparison with how good off Pebbles otherwise would have been, and secondly a comparison between her fate and the fates of other relevant people or with what kind of life it would reasonably have been expected that Wilma's child should have. Since Pebble's fate is significantly worse than these fates, Pebbles has on this second ground been exposed to an injustice in which Wilma is complicit. Wilma's act was thus on this second ground morally wrong.

A solution to the Non-Identity Problem should meet Boonin's demands for independence, robustness, and modesty (see introduction section). While I do believe that this theory could meet these requirements, at least after some further development, due to the lack of space, I have only briefly discussed this. Thus, it remains a question that others can take a stand on, and I can return to later.

University of Helsinki

References

Arrhenius, Gustaf, Ryberg, Jesper and Tännsjö, Torbjörn (2022). "The Repugnant Conclusion." In *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), edited by Edward N. Zalta & Uri Nodelman. URL = <https://plato.stanford.edu/archives/win2022/entries/repugnant-conclusion/>.

Axelsen, David V., Bidadanure, Juliana and Meijers, Tim, eds. (2021). *Luck Egalitarianism*. Routledge: New York.

Boonin, David (2014): *The non-identity problem and the ethics of future people*. Oxford, England: Oxford University Press.

Boonin, David (2019). Solving the Non-Identity Problem: A Reply to Gardner, Kumar, Malek, Mulgan, Roberts and Wasserman. *Law, Ethics and Philosophy,* 7, 127–156.

Gardner, Molly (2019). When Good Things Happen to Harmed People. *Ethical Theory and Moral Practice*, 22 (4), 893–908.

Elkins, David (2007). "Responding to Rawls: Toward a Consistent and Supportable Theory of Distributive Justice." BYU *Journal of Public Law*, 21, 267–322.

Huseby, Robert (2012). "Suffiency and Population Ethics." *Ethical Perspectives* 19 (2): 187–206.

McMahan, Jeff (2013). "Causing People to Exist and Saving People's lives." *The Journal of Ethics*, 17 (1/2), 5–35.

Lippert-Rasmussen, Kasper (2023). "Justice and Bad Luck." In *The Stanford Encyclopedia of Philosophy* (Spring 2023 Edition), edited by Edward N. Zalta & Uri Nodelman. URL = <https://plato.stanford.edu/archives/spr2023/entries/justice-bad-luck/>.

Lippert-Rasmussen, Kasper (2021). "Precis of Luck Egalitarianism." In: *Luck Egalitarianism*, edited by David V. Axelsen, Juliana Bidadanure and Tim Meijers.

Masny, Michal (2020). "On Parfit's Wide Dual Personal Principle." *The Philosophical Quarterly*, 70 (278), 114–139.

Parfit, Derek (2017). "Future People, the Non-Identity Problem, and Person-Affecting Principles." *Philosophy and Public Affairs* 45, 2.

Parfit, Derek (1984). *Reasons and Persons*. Oxford, England: Oxford University Press.

Parfit, Derek (1983). "Energy Policy and the Further Future: The Identity Problem." In *Energy and the Future*, Totowa, edited by MacLean D. and Brown P. N.J.: Rowman & Allanheld.

Piazza, Jared, Sousa, Paulo, Rottman, Joshua and Syropoulos, Stylianos (2018). "Which Appraisals are foundational to moral judgment? Harm, injustice, and beyond." *Social Psychological and Personality Science*, 1–11.

Thompson, Janna (2016). "Obligations of Justice and the Interests of the Dead." *International Journal of Applied Philosophy,* 30 (2): 289–300.

Zwolinski, Matt, Benjamin Ferguson, and Alan Wertheimer (2022). "Exploitation." In *The Stanford Encyclopedia of Philosophy* (Winter 2022 Edition), edited by Edward N. Zalta & Uri Nodelman. URL = <https://plato.stanford.edu/archives/win2022/entries/exploitation/>.

The future of political institutions - Foucault, genealogical critique, and the normative implications of his analysis of the state

TUUKKA BRUNILA

1. Introduction

The political and normative entailments of Michel Foucault's theory and his philosophical method have remained an issue for interpreting the meaning of his work. This is partly because Foucault himself is somewhat reluctant to define his own political orientation (Foucault 1994g, 593). However, more relevant is his explicit refusal to develop a normative framework for his method of analyzing the rationalities and practices of power. According to Foucault, his critical research is supposed to be immanent rather than transcendental, and its main concern is making visible the rationalities and conceptualizations operating in power relations and practices that seem to us as if they were self-evident or invisible (Foucault 1994d, 180). In this context, a normative critique is inappropriate as it would have to begin with principles that transcend the objects of analysis. Furthermore, normative criteria would also contradict Foucault's basic idea that there are no universal structures or immutable conditions—an idea that is the hypothesis behind his analyzes as their objective is to reveal the radically historically determined and politically changeable aspects of our present (Foucault 1994h, 574).

However, many have argued that Foucault's genealogical method is in fact contradictory as it both seeks to refrain from passing moral judgement, and yet it seems to imply that there is something about the objects of analysis that need to be struggled against. For example, in his explicit remarks re-

gards regarding genealogy, Foucault claims that genealogy is meant to support struggles against specific practices of power (Foucault 1997, 11–12). For this reason, political philosophers like Nancy Fraser, Jürgen Habermas, and Charles Taylor, argue that Foucault is either not being open about his normative stance, making his method contradictory, or his lack of a normative framework makes his method unfit for supporting struggles against domination (Fraser 1981; 1985; Habermas 1985; Taylor 1984). Furthermore, Mitchell Dean and Daniel Zamora claim that Foucault's methodology is normative as it focuses on the practices of governing rather than state power because it therefore serves to redirect political struggle from the state to the non-state practices of power (Dean and Zamora, 2021, 5). While they are critical of this shift of focus, others argue that this is an important development, as Foucault makes visible the sites of struggle that are not state-centered (see de Lagasnerie 2020, 74; Laval 2015, 30; cf. Brunila 2023).

In this article, I seek to re-evaluate Foucault's normative standpoint to assess the usefulness of "genealogical critique," as Martin Saar calls it (Saar 2007). In order to develop a normative reading of Foucault's genealogical critique, I focus on its relevance for the critique of political institutions. While many, Saar included, underscore genealogy's capacity to critically examine the production of subjectivity (see Oksala 2016), genealogy has been applied to institutions as well (e.g. Lichtenstein 2020). Furthermore, Scholars such as Karsten Schubert have argued that Foucault's methodological anti-universalism and its thesis that politics are inherently contingent lay the foundation for radical democratic institutions (Schubert 2021, 55). As I have argued elsewhere, such "post-foundational" political theories are normative in so far as they defend plural and open democratic institutions (Brunila 2022a; see Marchart 2018a). However, in order for post-foundationalist attempts to harness Foucault's methodological insight for normative reasons, the problem regarding his genealogical method's normativity will have to be answered.

While Schubert and others have developed their own Foucauldian universalist notions of freedom (Schubert 2019, see Mascaretti 2019), thus defending a universalist position that would create the basis for distinguishing between good and bad forms of governing, I focus on the way Foucault analyzes

political institutions and what normative ideas can be inferred from it. This way, one does not have to connect Foucault's work with a universalist position but, instead, I develop an approach to critiquing political institutions with normative entailments. I illuminate Foucault's critical method as one that underlines that political institutions are historically contingent, as they can be transformed for the means of various political rationalities and accommodate different practices of power. In this way, Foucault's idea is not to negate or dismiss these institutions, but simply to understand them as being inherently historically contingent. I argue that this is because Foucault's genealogy is meant to excavate rationalities and practices rather than institutions. Unlike Dean and Zamora argue, I claim that genealogy's ability to disclose the contingent nature of political institutions is what makes it possible to understand these institutions as being transformable and therefore a site of political struggle. In contrast, if one were to approach political institutions as having an immutable essence, the normative outcome would be to either affirm them as such or overthrow them.

A good counterexample for the kind of attitude that essentializes political institutions is Giorgio Agemben's political theory, which used Foucault's work to examine sovereign power (Agamben 1998), and whose main idea has been that the state and law are principally institutions of foundational sovereign violence. Recently, Agamben's work has become relevant both because of the states of exceptions and the alleged use of sovereign power throughout the world during the COVID-19 pandemic (cf. Lehtinen and Brunila 2021) and because of his own comments on these events (Agamben 2021). Relying on his unique Foucauldian reading of sovereignty and biopolitics, he has made controversial statements regarding pandemic governance by making comparisons between the Italian pandemic governance and National socialist rule (Agamben 2021, 36–37). Furthermore, Agamben claims that the pandemic makes us aware of the worrying issue of whether "some of the words that we keep on using—such as democracy, legislative power, elections, constitution—actually lost their original meaning a long time ago" (*ibid.*, 65). For Agamben, all these institutions have ultimately lost their original purpose due to the government's constant use

of exceptional measures, a practice that ultimately makes the state of exception the normal situation indistinguishable.

Agamben's theory seeks to interpret political institutions as having a core essence that transcends historical changes, meaning that these institutions are either good or bad, and they are politically useful only if they still retain this original essence. My argument is that this is contrary to Foucault's own method. The genealogist opens up possibilities for the future by trying to understand the struggles and developments taking place within specific institutions and how different practices of power can operate through them.

I develop my normative reading of genealogy by emphasizing its temporal nature. Instead of negating the past in a grand critical gesture by deeming it as something inherently corrupted, Foucault's analyzes can be seen as a way to interpret the past's transformative potential. This is because Foucault sees political institutions as being the result of historical processes, rather than the locus of universal essences, which one needs to either condemn or accept as such. Furthermore, the genealogical approach to institutions is normatively relevant as it seeks to change the way we understand how we would seek to transform the present.[1] In order to develop a normative reading of a critique of this kind, I focus on the principles of genealogical critique as a manner or attitude that approaches power from a very specific standpoint, which is temporal by nature. Genealogical critique brings together the past, the present, and the future for transformative reasons. It is interested in the past for the sake of present political struggles, the conditions of which have been shaped by past practices, power relations, and rationalities. By means of making present power relations transparent, the genealogist takes part in their own political context and for the sake of opening up new possibilities.[2]

---

[1] For Martin Saar, genealogical critique is mainly a manner of criticizing the present by means of analyzing its historical roots in the past. By making the historical origins of our political present intelligible, genealogical critique "opens up a space of possibilities for other attitudes, actions and identities" (Saar 2007, 294).

[2] This means that genealogy wants, as Johanna Oksala puts it, to "open our eyes to the need for a political criticism" (Oksala 2007, 88; Saar 2007, 318; see Tiisala 2017, 14).

2. Normative principles of genealogical critique

In this section, I discuss how Foucault understands genealogy. Instead of focusing on a methodological discussion regarding the details and possible inconsistencies of Foucault's method in analyzing the rationalizations of power (see Biebricher 2008), I want to clarify the central principles of genealogy to elaborate its normative entailments for critique of the political institutions. However, one should always keep in mind that genealogy is not a master-key to understanding Foucault's methodology. First of all, some scholars argue that Foucault is in fact not very concise with his method of analysis, even when it comes to genealogy itself (see Rehmann 2016). Second, Foucault distinguishes between "archeology", a method common in Foucault's early work, and genealogy. The former is meant as an excavation of discourses and their rules, practices, conjunctures, and ruptures (Foucault 1969, 188–190). While archeology is about making discourses and their principles explicit in order to re-discover the struggles and exclusions that take place in them, genealogy is a "tactic" to take part in these struggles (Foucault 1997, 11–12). Genealogy, therefore, means re-evaluating and participating in the struggles that are inherent knowledge production and practices of power.

In descriptive statements regarding genealogy, Foucault indeed understands genealogy as a normative endeavor. As Foucault puts it, archeology analyzes "the discourses that articulate how we think, talk and act," and genealogy "discloses the contingency of that which constitutes what we are *in order to not be like that anymore*" (Foucault 1994h, 574; emphasis added). There are two important ideas presented in this quote. Both are temporal, as they underline the role of genealogy for the present and the future. First, that genealogy reveals the *contingency* of those practices that constitute our subjectivity, that is, the way we currently exist and act. Our determinate existence is historical as it has been produced by various power-effects and discourses, which in themselves are not universal but local and particular to a specific time. Second, the discovery of this contingency leads to the possibility of transforming the way we exist and the powers that

format our present. This section focuses on elaborating these two ideas for the sake of critiquing political institutions.

However, some scholars have not been convinced about the normative side of Foucault's genealogical method. A notable example is Habermas, who accuses Foucault of "crypto-normativity", that is, of the fact that genealogical analysis smuggles its normative entailments into the analysis implicitly (Habermas 1985, 331). What Habermas means by this claim is that Foucault's method is contradictory. He makes this accusation by pointing out that, first, Foucault claims both that his analysis does not make normative statements regarding his object of analysis and that it does not constitute a scientific discourse capable of being objective (*ibid.*, 327). Genealogy, detached from normativity and objectivity, neither seeks to judge historical processes from a moral standpoint nor establish a science that would dominate over other forms of knowledge. Such a standpoint, for Habermas, is relativist as it cannot evaluate or disqualify normative and scientific claims (*ibid.*, 330–331). Foucault's standpoint is therefore incapable of giving reasons whether one should resist the practices and rationalities that he has made intelligible.

However, Foucault's analyzes seem to at least suggest that the effects of these practices are something to be struggled against.[3] In fact, he claims that the motives behind his research are political in so far as its objective is to emancipate us from the prevailing forms of exercising power (Foucault 1994d, 180). Therefore, it seems that Foucault simply masks his normative stance behind the veil of his explicit notions regarding his genealogical method. For Habermas, this shows that Foucault's method is contradictory and therefore useless for a critique of power.

Fraser argues that, because Foucault's theory of power makes it impossible to pass moral judgment, his work ends up "inviting questions which it is structurally unequipped to answer" (Fraser 1981, 281). Namely, it cannot establish a normative framework, which could distinguish between the good and bad practices of power. Similarly, Taylor argues that "Foucault's analyzes seem to bring evils to light; and yet

[3] Todd May describes this as a "genealogical ethos" that is apparent to Foucault's style of writing (May 2017, 170–171).

he wants to distance himself from the suggestion that would seem inescapably to follow that the negation or overcoming of these evils promotes a good" (Taylor 1984, 152). For both of them, Foucault cannot openly state what is wrong with the various practices of power that have been illuminated.

These critiques have resulted in numerous rebuttals. For example, Giovanni Maria Mascaretti and Daniele Lorenzini both argue that neither Habermas nor Fraser appreciate the specific way Foucault understands critique (Mascaretti 2019, 30; Lorenzini 2020).[4] However, unlike Mascaretti, I do not focus on Foucault's own ethical project of self-transformation (see Mascaretti 2019, 41–42), – a project that has to do with developing a universalist position regarding freedom (see Schuber 2021).[5] While I agree that Foucault indeed sees self-transformation as a way of realizing one's freedom,[6] I emphasize that, in the context of political institutions,[7] Foucault's notion of genealogy itself entails a normative manner of analyzing as it makes institutions intelligible in a way that directs political action.

---

[4] As Raymond Geuss points out, Foucauldian critique is not meant to be a manner of either affirming or negating in the sense of "X is bad," but a method for emphasizing the inherent dangers in X (Geuss 2002, 279).

[5] Here, I find Mattias Lehtinen's position regarding democratic institutions more plausible, as he develops an explicit normative framework from the ground up (see Lehtinen 2023). In contrast, Foucault's principles are too weak to be developed into a full-grown democratic theory.

[6] In this context, I Fraser points out that Foucault's understanding of freedom is too vague as it does not offer any idea what emancipation would actually mean. Rather, Foucault's lack of normative theory of freedom means that all he can offer is rejecting the current situation for "an unknown X", which he refuses to elaborate as a matter of principle (Fraser 1985, 180). Others, more sympathetic, have argued that Foucault's understanding of freedom is negative in so far as Foucault values freedom from being governed, rather than freedom to govern (see Brännström 2011, 124).

[7] In this way, I intend to evade Sandrine Rui's correct observation that it is rather peculiar that Foucault's concepts, which were originally meant to serve the purposes of critical inquiry, have been appropriated for a positive democratic programme (Rui 2013, 66). It seems to me that such a programme would need a more explicit notion of democracy, one that would go beyond Foucault's genealogical method.

I will discuss genealogy and political institutions in the next section and focus here on the two main temporal principles that are relevant regarding genealogy's normativity. These principles establish that genealogy is a method of revealing the contingency of our present and a participant in transforming it. Both derive from Nietzsche. He defines genealogy as a way of analyzing the past for the sake of understanding our present values and the powers that shape us in order to re-evaluate them. This task, according to Nietzsche, is only possible by means of looking at the historical conditions and their development (Nietzsche 2014, 253). *On the Genealogy of Morals* analyzes how the practices of punishment have produced the present moral values of his time (*ibid.*, 294–297).[8] The underlying hypothesis of this type of genealogy is that the history of our present is the history of struggle for power, and that all re-interpretation and transformation are the continuation of this struggle for power (*ibid.*, 313–314).

The basic principle regarding contingency, operative in Foucault's genealogy, is that the present is wholly contingent as there is nothing that transcends history. Institutions, identities, and practices of the present all have a historical origin. Contingency means that nothing in the present has risen out of necessity and therefore things could radically be otherwise. This type of research, as Thomas Marttila puts it, is based on "the ontological premise that social norms, values, beliefs, and rationales cannot reflect any external necessities such as the teleological course of history, objective material constraints, the inherent nature of human being, or the like" (Marttila 2015, 33). Ultimately, contingency entails that the foundation of our social relations and political order is determinate in the sense that it excludes other possibilities. This exclusion requires power to bring about and uphold this order against contestations.[9]

Contingency, therefore, implies the ever-present possibility of conflict in our present circumstances and the capacity to

[8] Similarly, Foucault argues that his motivation behind analyzing the prison system was a genealogy of morality (Foucault 1994i, 21).

[9] For Nietzsche, history does not progress towards a specific end but, instead, the movement of history is simply a series of struggles for greater power (Nietzsche 2014, 314)

contest it (see Marchart 2018b, 33). For Foucault, contingency refers to the fact that our social and political situation is born out of struggle. The task of genealogy is to bring about the decisions and exclusions that are part of the history of our own situation. Genealogy re-discovers the "discontinued, disqualified, non-legitimized" aspects that have been forcefully excluded in the past to make way for the specific practices and production of knowledge (Foucault 1997, 10).

Because our present social and political world is historically contingent, it means that the prevailing order is contestable. For genealogy, therefore, the task is to emancipate the subjugated (Foucault 1997, 11). This emancipatory potential is inherent to genealogy as it is meant to open up possibilities for change. "I do not conduct analyzes to say: here is how things are, you are trapped. I talk about these things only because I consider them to be transformable" (Foucault 1994c, 93). This means that the genealogist is not above the analysis, as if the motivations and methodological decisions would transcend the analysis of the past. Instead, Foucault sees the genealogy as tied to the present and enmeshed in power relations, meaning that it is not possible to remain in an objective position from which to evaluate various political systems (Biebricher 2008, 366; see Marchart, 2007, 4).

However, does this mean that Foucault is tied to the Nietzschean idea that genealogy, too, strives for power and that Foucault seeks to re-interpret the past to advance his own will to power? For Nietzsche, everything that takes place in history is simply striving for power, "and all seizing and rising to power is re-interpretation" (Nietzsche 2014, 314). To be sure, others inspired by Nietzsche argue that all politics is simply about striving for exerting one's will over others. For example, according to Max Weber, "who is taking part in politics, strives for power" (Weber 1992, 158–159). However, Foucault's point is not to replace one regime with another. Rather, the task of critique is to challenge, limit, transform, escape and displace present forms of governing for the sake of emancipation from extensive forms of governing. As Foucault puts it, such a task can be characterized as "the art of not being governed so much" (Foucault 2015, 37). While power itself might be something ineradicable, the way we are being governed can be made less forceful and total by means

of criticizing practices and opening up sites of resistance.[10] As I pointed out above, Foucault's idea is that genealogy would help us with emancipating ourselves from the prevailing practices of power.

When it comes to emancipation, which is made possible by the genealogical method, however, it does not entail an outright dismissal of past practices. As Foucault puts it in an interview, the point of genealogy is not to portray all power as bad but to understand the dangerousness inherent to its practices. This means that, rather succumbing to apathy, "if everything is dangerous, we can always do something" (Foucault 1994a, 386). Instead of diagnosing all endeavors to transform power-relations as doomed to fail, since all power is bad anyway, Foucault wants to show that the prevailing practices that we face currently are dangerous and that they should be resisted and transformed. However, this also means that change for the mere sake of change can be dangerous. As Foucault puts it, some of the most dangerous political traditions originate with the idea that everything in our social world needs to be transformed (Foucault 1994h, 575).

Based on Foucault's explicit remarks, genealogy is explicitly normative, as it is meant to take part in the struggle for the sake of opening up possibilities. This is evident in the temporal nature of genealogy, as it seeks to uncover the past for transforming the present and opening up future possibilities. Specifically, genealogy does not seek to negate the past but, instead, interpret it as dangerous and therefore transformable. In the next section, I argue that this insight is tied to Foucault's genealogical method as it is meant to excavate the various powers that affect us in the present rather than the history of various institutions and their origin. Instead of looking for Foucault's political convictions or values, like Dean and Zamora have done (2021), the method itself is normative[11] as it uncovers political institutions in a way that sees

[10] As Joonas Martikainen puts it, "the role of critical theory is to identify the ways the current social arrangements are oppressive towards a large majority of mankind" (Martikainen 2021, 13).

[11] Or according to Porcher, it is normative rather than normativist (Porcher, 2023).

them as transformable rather than as things to be overthrown.

3. Genealogical method in analyzing the state

I have now established the central principles of genealogical critique, which is a strategy to uncover the contingent and contestable historical basis of our present. According to Laura Jenkins, genealogy therefore analyzes the processes of "forming necessities, permanence, immobility, closure, and fatalism and concealing/negating or removing contingency" (Jenkins 2011, 160). In this section, I develop the normative entailments in Foucault's method by focusing on political institutions, specifically the state and law. I argue that, since the temporal character of genealogy is present here, as Foucault seeks to argue that the state and law are historically contingent and therefore transformable, this means that the way Foucault analyzes political institutions is also normative. The way Foucault approaches these institutions is by pointing out that throughout history they have been appropriated by various political rationalities and practices of power. These forms of rationalizing power, such as biopolitics, and the practices that they entail, such as biopower, are distinct from the institutions through which they operate. The state, instead of forming its own distinct essence, is a historically contingent conglomerate of various prevalent rationalities and practices. I argue that it is this approach to political institutions that makes Foucault's genealogy normative.

In the context of state and law, Foucault was very open about the normative entailments of his work. For him, political theory is stuck on analyzing power as sovereign power. "We are still attached to a specific image of power as law, power as sovereignty, of which analyzes of legal theory and monarchical institutions draw upon. It is especially from this image, that is, privileging the theory of law and sovereignty, from which we have to free ourselves in order to analyze power in its concrete workings and historical processes." (Foucault 1976 118–119.) Here, Foucault's temporal and therefore normative ethos is fairly visible as he argues that the way we understand power is tied to emancipation.

Foucault claims that the state is "heartless" (Foucault 2004a, 79). By this, he simply means that the state should not be analyzed as having its own essence but rather as a target of various discourses and the collection of multiple practices (Foucault, 2004a 5–6; Foucault 1994b, 150). This is directly tied to his methodological principles regarding analysis: instead of beginning with universals such as the state, Foucault wants to study various powers that expand throughout society (Foucault 2004a, 5). Instead of assuming the state as the centralisation of power, Foucauldian genealogy begins with the practices of power in order to understand the historical nature of state power and its process of centralizing various capacities. Multiple practices become colonized by the state to serve certain interests, such as bourgeoisie ones (Foucault 1997, 29–30). In his lecture series on liberalism, *The Birth of Biopolitics,* Foucault describes this process of colonization as "statification" (*étatisation*), which namely means bringing various practices and powers under state control (Foucault, 2004a, 79). The state is a superstructure that is a historically contingent collection of various practices and powers, or, the site of meta-power (Foucault 1994b, 150).

Foucault's main methodological idea regarding why the state should not be the starting point for political theory and analysis is that this will lose sight of powers and its practices that operate throughout social relations. For this reason, Foucault did not find fruitful the distinction between state and society, that is, the distinction between power and the social. We are not simply being influenced from above or by the state. Instead, "it should rather be assumed that multiple relations of forces operate in the apparatuses of production, family, minor groups and institutions are the basis for large divisions that traverse throughout the social body" (Foucault 1976, 124). All social relations implicate power relations; and they cannot be reduced to the top-down power of the state.

In order to analyze how power succeeds in permeating the individual and social relations, one must forgo the idea of power as mere repression and prohibition (Foucault, 1976, 20). Power as repression, Foucault claims, is based on a juridical notion of power as simply establishing external limits to the citizen's conduct. In this way, power-relation is simply a negative relation that establishes a rule that prohibits (Fou-

cault 1976, 110–111). Furthermore, such a conception understands power as a unity, which means that power functions always in the same way by means of law and prohibition "from the state to the family, from prince to the father, from the tribunal to quotidian small punishment, from the social domination to the constructive structures of the subject itself" (Foucault 1976, 112). Such a notion of power is incapable of genealogically revealing the different and multifaceted nature of these various power relations. Specifically, Foucault warns against beginning an analysis of modern power relations with an abstract idea of law or the state, as that will simply serve to obfuscate power.

In order to argue that the sovereign model has become obsolete, Foucault analyzes the medieval model of state power, or the "royal model", which identifies the state with the wielder of sovereign power. Ever since the medieval times, western societies have understood power by means of law (Foucault 1976, 115). This royal model understands power as being centralized and legitimated by means of law and subjection through citizenship. Here, power is strictly understood as prohibition, which is the will of the sovereign king exerting power over subjects (Foucault 1994b, 150).

Sovereign power is based on the idea that power is unified to the sovereign for the state to be capable of establishing subjection by means of overpowering its citizens (Foucault 1997, 38–39). This means, above all, coercion. Theories of sovereign power seek to make this coercion legitimate by means of arguing why such domination is in fact justified. Here, Hobbes is the paradigmatic example. Power must be unified to the sovereign in order to make sure that subjects limit their actions in fear of punishment (Hobbes 2018 XIII, § 62–63).[12] Without coercive power, Hobbes claims, peace within a society would be lost as "covenants, without the Sword, are but Words, of no strength to secure a man at all" (Hobbes 2018 XVII, § 85). The only way to uphold order, is to confer all

[12] Similarly, Rousseau claims in *The Social Contract* that "the sole means that they still have of preserving themselves is to create, by combination, a totality of forces sufficient to overcome the obstacles resisting them, to direct their operation by a single impulse, and make them act in unison" (2008 I, vi; emphasis added).

power to the sovereign who will "reduce all their Wills, by plurality of voices, unto one Will", that is, to form the sovereign is to submit will and judgment to the sovereign (Hobbes 2018 XVII, § 87; see Foucault 2013, 29).

Foucault understands the royal model as a theory that reduced power to law and law to coercion, so that the royal model thinks that "law is always referred to as a sword" (Foucault 1976, 189). Ultimately, it is simply the subject's preference to live rather than be killed that establishes legitimate subjection and absolute centralized coercive power (Foucault 1997, 82). Violence therefore defines the sovereign use of power since the fear of violence will make sure that citizens obey laws and limit their actions for the sake of civil order (Foucault 2013, 12).

Foucault's genealogies make analyzing the state in the old medieval model of equating it with sovereign power obsolete for various reasons. For example, the development of modern life sciences makes it possible for those who govern to target populations and individuals based rather than simply by means sovereign coercion (Foucault 1994e, 192–193). Foucault calls this kind of power "biopower," the power targeting the population as an object of medical and biological practices. It is distinct from the state's sovereign power as it seeks to maximize the health of the population or the individual by means of healthcare, statistics and other such practices (Foucault 1976, 181; Foucault 1997, 214; Oksala 2010, 36; Erlenbusch-Anderson 2020, 8). Similarly, in *Discipline and Punish*, Foucault analyzes how modern disciplinary practices and supervision target the individual's body to produce desired conduct (Foucault 2011, 161–163).

The development of these powers targeting the body and the population alter and transform the practices that states have available. For example, Foucault describes how punishment in the royal model has a juridical-political function to remind the citizens of the power of the monarch. For this reason, punishment was always public and was meant to restore the king's authority in the minds of subjects (Foucault, 2011 59–60). In contrast, new disciplinary methods target the individual's body to direct behavior (Foucault, 2011, 353). Similarly, in the *History of Sexuality*, Foucault argues that during early modernity governing shifts dealing with legitimate sub-

jection of citizens into focusing on populations to influence the population's birth-rate, life expectancy, fertility, health, and other aspects central to upholding vitality (Foucault 1976, 36).

For this reason, Foucault argues that neither the state nor law can be analyzed as historically transcendental categories. While the royal model understood law as the sovereign power's instrument for prohibition, the disciplinary methods transform law into a tool for the normalization and production of normal behavior (Foucault 2011, 355). Therefore, Foucault does not argue that law has become obsolete but that as a practice it has been transformed for new purposes and applications (Foucault 1976, 116). Instead of equating the state with sovereign power, Foucault argues that the state is a "meta-power" that appropriates various powers and practices that are available to it. As a meta-power, the state is a historical entity that takes over or "stratifies" historically determined powers. This centralization of powers is distinct from sovereign power as the latter is merely one way the state exerts its power. While Foucault agrees that sovereign power is still relevant – as powers targeting citizens, populations, and bodies *complement* one another rather than replace – the state is an altogether different entity from its medieval equivalent (Foucault 2004b, 109–110).

Inherent to these analyzes of various rationalizations and practices of powers is that law and state as institutions are historically contingent and are not defined by a specific essence. The critique regarding law in the *History of Sexuality* is not targeted against analyzes of law as such but against the royal model's interpretation of law as being outdated. Furthermore, in contrast to Dean and Zamora's argument that Foucault shifts our focus from the state, the critique of the royal model is not a critique of analyzes of the state. Instead, Foucault's genealogical analysis shows that identifying the state with sovereign power is not enough. What is needed is a conception of political institutions that is not based on establishing fixed essences, one that would understand the relationship between law and coercion, and state and sovereignty, as historically contingent.

I will now move onto further developing the normative aspect of the genealogical approach to political institutions.

As I mentioned in the introduction, others, too, have emphasized this. For example, Christian Laval argues that Foucault's notion of the multiplicity of power achieves multiplying the sites of resistance to non-legal and other social spheres (Laval 2015, 41–42; see Foucault 1976, 126–127). Similarly, Geoffroy de Lagesnerie has argued that a critique that "exalts the law, the political, or sovereignty is not only unsatisfactory but, on the contrary, potentially regressive and reactionary" (de Lagasnerie 2020, 74). However, I do not agree with these claims that Foucault's theory necessitates a complete dismissal of law as a site of struggle (see Laval 2015, 30). To be sure, in the royal model resistance to power takes place solely by means of law (Foucault 1976, 116). This meant specifically legal struggles to limit and restrict sovereign power and bind it to law (Foucault 1997, 31), so that law is the way in which the excessiveness and absoluteness of sovereign power is countered (Foucault 2004a, 8–10). Instead, as I have argued in this section, Foucault does not establish a general conception of law but understands it as an instrument that changes according to the practices of power and its relations (see Brännström, 2014; cf. Mazères 2017). It is this aspect that I want to underline in the next section.

4. Genealogy and the future after the pandemic

So far, I have established that genealogical critique is meant to analyze the past for the sake of understanding the political possibilities of the present. As Martin Saar puts it, genealogical critique is a method that approaches the present political situation through the analysis of its historical origins (Saar 2007, 9). Genealogical analysis clarifies the way in which power constitutes our own identity and political practices. In this sense, it is always an immanent critique of the political reality that we inhabit (Saar 2007, 21, 69, 222). By showing how these origins are in fact contingent, a genealogist opens up opportunities to transform our political future (Saar 2007, 294; Oksala, 2016, 7; Marchart, 2018b, 31-33; Lemke 2019, 374). I will elaborate why this approach, in the context of political institutions, is in fact normative as it entails a specific analysis of our political present.

Here, I seek to problematize Agamben's reading of what can be inferred from Foucault's genealogical analyzes of state power. In contrast to Agamben, Foucault wanted to counter reproducing what he called the "royal model," that is, the monarchist image of power as being centralized to the king. For this reason, Foucault engaged in the genealogy of the state as the history of various practices that are more or less centralized to the state. This historical account is not about a temporary succession of exclusionary forms of power, but simply their correlation and the way in which they function together. New forms of power do not simply replace sovereignty and judicial practices. Instead, as Oksala puts it, sovereign power is complemented by practices such as disciplinary and biopolitical practices (Oksala 2012, 94).

For Foucault, political theory seemed to be stuck on sovereign power, so much so that one of his main motivations was to "cut short the recurring summoning of the master and the monotone affirmation of such power" (Foucault 2004b, 56). By this, he meant that the fact that political theory was still stuck on the royal model was not only descriptively false, but also normatively problematic, as it tends to affirm the existence and justification of such power as the sole form of political power. Foucault did not mince words when he stated that "we need a political philosophy that does not base itself on the problem of sovereignty, […], the head of the king needs to be cut, and this has not yet been done in political theory" (Foucault 1994b, 150). In analyzing political institutions, such as the state or law, Foucault's methodology amounts to a specific approach, one that underscores their historically contingent nature, rendering them dangerous rather than evil.

The main take-away is the critique of identifying the state with sovereign power. The main issue for Foucault is analytical: political theorists are incapable of recognizing the multifaceted and productive nature of power. In order to illuminate what is normative about this, I contrast it to a branch of political theory that is first of all still tied to the royal model of power and, second, has become problematic for our present purposes after the pandemic. I refer to theories of the state of exception, that is, both descriptive and normative theories regarding the state's emergency measures during abnormal situations. By means of this discussion, I

intend to flesh out what is normatively unique to genealogical critique.

Like genealogists, I will begin with a short discussion about the past theories of the state of exception by discussing Carl Schmitt's influential theory. For Schmitt sovereign power is the power capable of making decisions and upholding order. Schmitt infamously defines the sovereign as the one who decides on a state of exception (Schmitt 2015a, 13). The exceptional situation is a grave threat to the political order and must be countered with power capable of acting without legal limits.[13] Similar to Hobbes, Schmitt argues that such a power that can act extra-legally has to have the highest authority and a centralized power structure (Schmitt 2015a, 18). Furthermore, like Hobbes, this kind of coercive power is necessary for law to exist, because power creates order in which norms and laws can become valid as "there is no norm that would be valid in chaos" (Schmitt 2015a, 19). This means that, because the law is only valid insofar as there is an authority capable of upholding, enacting and interpreting it, sovereign decision is the foundation of law (Schmitt 2015a, 15).

Schmitt develops a constitutional theory of exceptional situations to authorize legally unbound coercive power. Since the decision a sovereign must make is an exception to the norm, this decision is not legal but political and therefore derives its legitimacy from pre-legal political legitimacy (Schmitt 2015b, 41). This political legitimacy is therefore the essence of state constitution rather than specific constitutional laws (Schmitt 1993 3–5, 148–149). In emergencies, the sovereign is not bound by laws but is authorized by the political order (Schmitt 1993, 26–27). It is for this reason that sovereign power can legitimately put the entirety of the constitution aside.

---

[13] According to Kim Lane Scheppele, the main principle of Schmitt's theory of exceptional measures is that "the sovereign must have all of the lesser-included powers-for example, the power to decide when the situation has ceased to be 'normal,' thereby justifying the declaration of emergency, the power to determine when the emergency is over so that the rule of law may be safely restored, and the power to specify which political actors normally protected by the rule of law lose their protection in the interim" (Scheppele 2004, 1010).

Schmitt's theory has been appropriated both for normative and descriptive reasons. A well-known normative appropriation is Eric A. Posner and Adrian Vermeule's defense of Schmittian principles after the September 11 attacks. In line with Schmitt, they argue that "ex ante legal rules cannot regulate crises in advance, because unanticipated events will invariably arise" (Posner and Vermeule 2011, 42). An emergency requires an exception to law, as the general nature of law is by definition incapable of anticipating exceptional events, meaning that there must be a sovereign authority capable of interpreting specific events as warranting emergency measures. In descriptive context, Agamben has argued that Schmitt is right that law cannot account for the exception and therefore sovereign power is a necessary aspect of any legal order (Agamben 1998, 21; Agamben 2005, 1–2). For this reason, Agamben argues, the legal order itself normalizes the exceptional measures, as its very validity is dependent on sovereign power.

During the pandemic, both of these perspectives have become relevant as political forces have sought to unbind sovereign power. These theories might be useful for understanding the dangers inherent in emergency measures. To a greater or a lesser extent, Hungary and France legislated during the pandemic in order to authorize the executive power with far-reaching extra-legal powers. Both enacted legislation that enabled the state of exception practically to continue for an indefinite amount of time (Basilien-Gainche 2022, 441–444; Kovács 2022, 262–263). Furthermore, one could argue in line with Agamben's idea that, at least in these countries, the exception has indeed become the norm (Agamben, 2021, 84), as both countries have also upheld a state of exception since 2015. In Hungary, the government decided on a "migration emergency" (Halmi, 2021, 304; Kovács 2022, 260). France declared a state of exception due to the November 2015 Paris terrorist attacks. Both of these states of exceptions are still in place (Basilien-Gainche 2022, 434).

While referring to Schmitt's understanding about exceptional measures is indeed relevant, it matters how we approach these measures. In contrast to Foucault, Agamben's analysis generalizes Schmitt's theory to concern all state measures in a way that seems to give ground to the worri-

some normative aspects of Schmitt's theory. This is especially the case in recent theories of exceptional measures before the pandemic. For example, Gian Giacomo Fusco argues that all exceptions are by nature Schmittian (Fusco 2020, 33). Like Posner and Vermeule, he seems to ascribe to the royal model of power arguing that the state is identical with sovereign power. On the one hand, if an analysis of exceptional measures starts with the royal model, exceptional measures seem necessarily Schmittian and therefore must be rejected outright. On the other hand, if exceptional measures were seen as necessary during the pandemic (see Pozen and Scheppele 2020), then this kind of theory has to accept Schmitt's normative theory.

Agamben is a notorious example here as his theory identifies sovereign power with political institutions (Agamben 1998, 6). He fails to underscore the contingent and historical aspects of Foucault's theory of the state. Instead of arguing that law and the state are contingent institutions that change over time, Agamben confounds them with sovereign power in a way that critics have described as transcendental. As Sergei Prozorov puts it, "to diagnose the present or any other state of affairs as a global state of exception, itself indistinct from the normal state and hence not really exceptional, is hardly a valuable insight, even when it might be accepted as a logical implication of the definition of the 'original activity' of sovereign power" (Prozorov 2021, 10). By identifying the law and the state with sovereign power, Agamben loses the sight of Foucault's ideas regarding the role of the critique and possibilities of resistance. If law is fundamentally bad, a mere instrument for sovereign violence, then the only outcome is the dismissal of law altogether. According to Leila Brännström, Agamben's "assumptions are built on a perception of law as a machine whose workings, effects, and possibilities are given beforehand – once and for all" (Brännström 2008, 23). In the context of the pandemic, one might argue against Agamben that he overlooks the fact that in many countries the law was able to limit sovereign power and uphold the rule of law.

Foucauldian genealogy does not deny the possibility that sovereign power could take over. As I pointed out above, for Foucault the role of critique is to underline the dangerousness

of various practices. However, simply because a political institution is dangerous does not mean that it should be dismissed categorically. Rather, genealogy opens up possibilities for resistance that might take forms that appropriate legal or state measures. For example, in contrast to Posner and Vermeule, who argue that the executive is *de facto* legally unbound (Posner and Vermeule 2011, 207–208), both judicial and legislative branch were vital to upholding the rule of law and administering legal limits to executive power (see Farzamfar and Salminen 2020; Scheinin and Molbæk-Steensig 2021). Similarly, in his political activism, Foucault himself argued for the rights of the governed against excessive governing (Foucault 1994j, 362; Foucault 1994f, 390). Furthermore, we should analyze exceptional measures without ascribing to the royal model in order to see that during the pandemic, the state operated on multiple levels as a meta-power. For example, the state distributed vaccinations, administered quarantines and controlled border-crossing. For Foucault, such pandemic measures are all distinct and cannot be reduced to sovereign power (Foucault 2004b, 61–63; Foucault 2011, 232–233). In this sense, the state needs to be analyzed as a meta-power that is contingent and therefore a possible object of future transformations.

5. Conclusions

In this article, I have argued that the principles operative in Foucault's genealogical analysis are normative as they affect the way in which genealogy approaches political institutions. I did this by identifying the temporality of Foucault's genealogy. Genealogy seeks to understand the political struggles of the present by means of understanding their historical roots in the past in order to open up new political possibilities in the future. I clarified that the basic principles of genealogical critique are contingency and emancipation. First, genealogy establishes the contingency of the present practices of power and its relations. Second, genealogists engage in these studies to take part in emancipating the present from domination. After discussing these principles, I presented Foucault's genealogical approach to political institutions. I argued that for Foucault, instead of being universal and ahistorical, political

institutions are historically contingent and accommodate various forms of political rationalities and practices of power. For example, the genealogy of the state sees the state as a form of meta-power, which varies over time according to the power relations and practices of its time. From this discussion of the state as meta-power, I elaborated on how Foucault's approach to political institutions differs from contemporary political theories that instead seek to understand the state solely as wielder of sovereign power. I argued that, from the genealogist's perspective, such theorists are both descriptively and normatively problematic, especially after the pandemic, as they lose sight of other forms of power at play during exceptional situations such as a pandemic, but also because they present us with scant normative possibilities for transformation.

However, this does not necessarily mean that Foucauldian genealogy is the right approach to political institutions. According to Johanna Oksala, Foucault's approach to the state was meant to counter analyzes of the state as "the root of all political problems" by emphasizing that the state encompassed multiple various practices and rationalities (Oksala, 2012, 30). While this was indeed his intention, as I have pointed out above, perhaps the state is still something more than just meta-power. As some Marxists have argued, state power is an important factor in understanding the reproduction of the capitalist means of production (Bonefeld 2019; see Brunila 2022). This ties in with Fraser's critique that Foucault's refusal to discuss ideologies, economic domination, and state interests, coupled with his focus on micro-powers and the capillary character of modern power, simply serves to hide domination and the workings of the ruling classes (Fraser 1985, 280–281).[14] Similarly, because Foucault's analysis of governmentalities did not begin with analyzing the state (Foucault, 2004a, 4-5; Brännström, 2014, 42; Behrent,

[14] While Foucault does indeed point out that the bourgeoisie have indeed colonized various practices for their own interests (Foucault 1997, 29–30), he denies the idea that power is something that power could be seen as serving some subject's specific interests. Power is intentional as it serves specific goals and objectives, but these goals and interests are not the outcome of a decision (Foucault 1976, 124–125).

2019, 10), Dean and Zamora argue that it is specifically Foucault's method that has redirected contemporary political in the wrong direction with its disregard of the state and sovereign power (Dean and Zamora, 2021, 201).

My contribution elaborates the normatively specific way Foucault's genealogical method indeed discloses political institutions – and it therefore has a determinate scope with its limits. By separating the practices of power and political institutions, genealogy ultimately refuses to discuss state power as such. The state and other political institutions are a conglomerate of various practices. It becomes unclear if the state indeed can be anything whatsoever. Obviously the law and the state can function as instruments, but as instruments they might serve some interests and objectives better than others. Furthermore, when it comes to transforming the state, even if it is a historically contingent phenomenon, its possibilities aren't limitless.

University of Turku

Acknowledgements

This work was supported by the Strategic Research Council at the Academy of Finland (grant number 345950). The author would also like to thank anonymous reviewers for their insightful and helpful comments.

References

Agamben, G. (1998). *Homo Sacer.* Translated by D. Heller-Roazen, Stanford, Stanford University Press.

Agamben, G. (2005). *State of exception.* Translated by K. Attell, Chicago, University of Chicago Press.

Agamben, G. (2021). *Where are we now? The Epidemic as Politics.* Translated by V. Dani, Lanham, Rowman & Littlefield.

Basilien-Gainche, M.-L. (2022). "The French Management of COVID-19: Normalisation of Regimed of Exception and Degradation of the Rule of Law." In *Routledge Handbook of Law and the COVID-19 Pandemic*, edited by J. Grogan, & A. Donald. London: Routledge, 434–444.

Behrent, M. C. (2019). "A Liberal Despite Himself: Reflections on a Debate, Reappraisals of a Question." In *Foucault, Neoliberalism, and Beyond*, edited by Sawyer, S. W. & Steinmetz-Jenkins, D. London, Rowman & Littlefield, 1–2.

Biebricher, T. (2008). "Genealogy and governmentality." *Journal of the Philosophy of History*, 2 (3), 363–396.

Bonefeld, W. (2019). "Ordoliberalism, European Monetary Union and State Power." *Critical Sociology*, 45 (7–8), 995–1010.

Brunila, T. (2022a). *The autonomy of the political and the authority of the state: Carl Schmitt and the de-politicisation of the economy*. Helsinki: University of Helsinki.

Brunila, T. (2022b). "Depoliticization of politics and power: Mouffe and the conservative disposition in postfoundational political theory." *Frontiers in Political Science*, 4.

Brunila, T. (2023). Miten tutkia valtiovaltaa? [How to study stateauthority?] *Tiede & Edistys*, 49(2).

Brännström, L. (2008). "How I learned to stop worrying and use the legal argument: A critique of Giorgio Agamben's conception of law." *No Foundations,* 5, 22-49.

Brännström, L. (2011). "Den reflekterande olydnadens konst." *Fronesis,* 36–37, 116–125.

Brännström, L. (2014). "Law, Objectives of Government, and Regimes of Truth." *Foucault Studies,* 18, 173–194.

de Lagasnerie, G. (2020). *Foucault Against Neoliberalism?* Translated by M. MacLellan. Lanham: Rowman & Littlefield.

Dean, M. & Zamora, D. (2021). *The Last Man Takes LSD: Foucault and the End of Revolution.* London: Verso.

Erlenbusch-Anderson, V. (2020). "The Beginning of a Study of Biopower: Foucault's 1978 Lectures at the Collège de France." *Foucault Lectures,* III, 5–26.

Farzamfar, M. & Salminen, J. (2020). "The Supervision of Legality by the Finnish Parliamentary Ombudsman during the COVID-19 Pandemic." *Nordisk Administrativt Tidsskrift,* 99, 1–19.

Foucault, M. (1969). *L'archéologie du savoir.* Paris: Gallimard.

Foucault, M. (1976). *Histoire de la sexualité: La volonté de savoir.* Paris: Gallimard.

Foucault, M. (1994a). "À propos de la généalogie de l'éthique: un aperçu du travail en cours." In *Dits et Écrits 1954–1988, IV, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard, 383–411.

Foucault, M. (1994b). "Entretien avec Michel Foucault", in *Dits et Écrits 1954–1988, III, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 140–160.

Foucault, M. (1994c). "Entretien avec Michel Foucault". in *Dits et Écrits 1954–1988, IV, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 41–95.

Foucault, M. (1994d). "Est-il donc important de penser", in *Dits et Écrits 1954–1988, IV, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 178–182.

Foucault, M. (1994e). "Les mailles du pouvoir", in *Dits et Écrits 1954–1988, IV, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 182–201.

Foucault, M. (1994f). "Lettre à quelques leaders de la gauche", in *Dits et Écrits 1954–1988, III, 1988–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 388–390.

Foucault, M. (1994g). "Polémique, politique et problématisations", in *Dits et Écrits 1954–1988, IV, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 178–182.

Foucault, M. (1994h). "Qu'est-ce que les Lumières?", in *Dits et Écrits 1954–1988, IV, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 662–578.

Foucault, M. (1994i). "Table ronde du 20 mai 1978", in *Dits et Écrits 1954–1988, IV, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 20–34.

Foucault, M. (1994j). "Va-t-on extrader Klaus Croissant?", in *Dits et Écrits 1954–1988, III, 1980–1988*, edited by D. Defert, F. Ewald, F. & J. Lagrange. Paris: Gallimard**,** 361–365.

Foucault, M. (1997). *Il faut défendre la société: cours au Collège de France, 1975–1976.* Paris: Gallimard/Seuil.

Foucault, M. (2004a). *Naissance de la biopolitique: cours au Collège de France, 1978–1979.* Paris: Gallimard / Seuil.

Foucault, M. (2004b). *Sécurité, Territoire, Population.* Paris: Gallimard/Seuil.

Foucault, M. (2011). *Surveiller et punir. Naissance de la prison*. Editions Gallimard.

Foucault, M. (2013). *La Société Punitive.* Paris: Gallimard/Seuil.

Fraser, N. (1981). "Foucault on modern power: Empirical insights and normative confusions." *Praxis international*, 1(3), 272–287.

Fraser, N. (1985). "Michel Foucault: A' young conservative'?" *Ethics*, 96(1), 165–184.

Foucault, M. (2015). *Qu'est-ce que la critique? Suivi de La culture de soi.* Paris: Vrin.

Fusco, G. G. (2020). "Exception, fiction, performativity." In *States of Exception: Law, History, Theory*, edited by C. Cercel, G. G. Fusco & S. Lavis. London: Routledge, 15–33.

Geuss, R. (2002). "Kritik, Aufklärung, Genealogie." *Deutsche Zeitschrift für Philosophie*, 50(2), 273–282.

Habermas, J. (1985). *Der philosophische Diskurs der Moderne, Zwölf Vorlesungen,* Frankfurt: Suhrkamp.

Halmai, G. (2021). "The Pandemic and Illiberal Constitutional Theories", in *Pandemocracy in Europe: power, parliaments and people in times of COVID-19*, edited by M. C. Kettemann & K. Lachmayer. London: Hart Publishing, 229–314.

Hobbes, T. (2018). *Leviathan.* Cambridge, Cambridge University Press.

Jenkins, L. (2011). "The Difference Genealogy Makes: Strategies for Politicisation or How to Extend Capacities for Autonomy." *Political Studies,* 59, 156–174.

Kovács, K. (2022). "The Covid-19 Pandemic: A Pretext for Expanding Power in Hungary." In *Routledge Handbook of Law and the COVID-19 Pandemic,* edited by J. Grogan, & A. Donald. London, Routledge, 259–270.

Laval, C. (2015). "La Productivité du pouvoir." In *Marx & Foucault V: Lectures, usages, confrontations*, edited by C. Laval, L. Parltrinieri & FL. Taylan. Paris: La Découverte, 29–44.

Lehtinen, M. and Brunila, T., (2021). "A Political Ontology of the Pandemic: Sovereign Power and the Management of Affects through the Political Ontology of War." *Frontiers in Political Science*, 3.

Lehtinen, M. (2023). *Political Ontology and Imagination: Contours of a Democratic Imaginary*. Helsinki: University of Helsinki.

Lemke, T. (2019). *Foucault's Analysis of Modern Governmentality: A Critique of Political Reason.* London: Verso Books.

Lichtenstein, E.B. (2020). "On the Ways of Writing the History of the State." *Foucault Studies*, 28, 71–95.

Lorenzini, D. (2020). "On possibilising genealogy." *Inquiry*, 1–22.

Marchart, O. (2007). *Post-foundational political thought: political difference in Nancy, Lefort, Badiou and Laclau.* Edinburgh: Edinburgh University Press.

Marchart, O. (2018a). *Thinking antagonism: Political ontology after Laclau*. Edinburgh: Edinburgh University Press.

Marchart, O. (2018b). *Das Unmögliche Object - Eine Postundamentalische Theorie der Gesellschaft.* Berlin: Suhrkamp.

Martikainen, J. (2021). *Political Poverty as the Loss of Experiental Freedom.* Helsinki: Faculty of Social Sciences.

Marttila, T. (2015). *Post-foundational discourse analysis: from political difference to empirical research.* Houndmills, Basingstoke, Hampshire: Palgrave Macmillan.

Mascaretti, G. (2019). "Foucault, normativity, and freedom: A reappraisal." *Foucault Studies*, 23–47.

May, T., (2014). "Governmentality." In *The Cambridge Foucault Lexicon*, edited by L. Lawlor & J. Nale. Cambridge, Cambridge University Press, 175–181.

Mazères, J.A. (2017). "Normativité, vérité, gouvernementalité: figures du juridique chez Michel Foucault." *Revue interdisciplinaire d'études juridiques*, 79 (2), 55–75.

Munro, V.E. (2003). "On power and domination: feminism and the final Foucault." *European Journal of Political Theory*, 2 (1), 79–99.

Nietzsche, F. W. (2014). *Zur Genealogie der Moral.* Berlin: de Gruyter.

Oksala, J. (2007). *How to read Foucault.* London: Granta Books.

Oksala, J. (2010). "Violence and the Biopolitics of Modernity." *Foucault Studies,* 10**,** 22–43.

Oksala, J. (2012). *Foucault, politics, and violence.* Evanston: Northwestern University Press.

Oksala, J. (2016). *Feminist Experiences: Foucauldian and phenomenological investigations.* Evanston: Northwestern University Press.

Porcher, F., (2023). "Critique immanente et critique généalogique. Le problème des normes de la critique sociale." *Strathèse*, 7.

Posner, E. A. & Vermeule, A. (2011). *The Executive Unbound: After the Madisonian Republic.* Oxford: Oxford University Press.

Pozen, D. E. & Scheppele, K. L. (2020). "Executive underreach, in pandemics and otherwise." *American Journal of International Law,* 114**,** 608–617.

Prozorov, S. (2021). "A Farewell to Homo Sacer? Sovereign Power and Bare Life in Agamben's Coronavirus Commentary." *Law and Critique,* 34**,** 68–80.

Rehmann, J. (2016). "The Unfulfilled Promises of the Late Foucault and Foucauldian 'Governmentality Studies'." In *Foucault and Neoliberalism*, edited by D. Zamora & M. C. Behrent. Cambridge: Polity press, 134–158.

Rousseau, J.-J. (2008). *The Social Contract.* Translated by C. Betts. Oxford: Oxford University Press.

Rui, S., (2013). "«Où donc est le danger?» Participation et usages de Foucault." *Participations*, 2, 65–86.

Saar, M. (2007). *Genealogie als Kritik: Geschichte und Theorie des Subjekts nach Nietzsche und Foucault.* Frankfurt: Campus Verlag.

Scheinin, M. & Molbæk-Steensig, H. (2021). "Human Rights-Based Versus Populist Responses to the Pandemic." In *COVID-19 and Human Rights*, edited by M. Kjaerum, M. F. Davis & A. Lyons. London: Routledge, 19–34.

Scheppele, K. L. (2004). "Law in a Time of Emergency: States of Exception and the Temptations of 9/11." *University of Pennsylvania Journal of Constitutional Law,* 6, 1001–1083.

Schmitt, C. (1993), *Verfassungslehre.* Berlin: Duncker & Humblot.

Schmitt, C. (2015a). *Politische Theologie.* Berlin: Duncker & Humblot.

Schmitt, C. (2015b). *Der Begriff des Politischen: Text von 1932 mit einem Vorwort und drei Corollarien.* Berlin: Duncker & Humblot.

Schubert, K., (2020). "Freiheit und Institution. Für eine anti-anarchistische Foucault-Lektüre." *ZPTh–Zeitschrift für Politische Theorie*, 10(1), 15–16.

Schubert, K. (2021). "Der letzte Universalismus: Foucaults Freiheitsdenken und die Begründung von radikaler Demokratie im Postfundamentalismus." In *Das Politische (in) der Politischen Theorie*, edited by O. Flügel-Martinsen, F. Martinsen & M. Saar. Baden-Baden: Nomos Verlag, 43–48.

Taylor, C., (1984). "Foucault on Freedom and Truth." *Political Theory*, 12(2), 152–183.

Tiisala, T., (2017). "Overcoming 'the present limits of the necessary': Foucault's conception of a critique," *The Southern Journal of Philosophy*, 55, 7–24.

Weber, M. (1992). *Gesamtausgabe. Abt. 1, Schriften und Reden, Bd. 17: Wissenschaft als Beruf: 1917/1919: Politik als Beruf; 1919.* Tubingen: Mohr.