BALKAN JOURNAL OF PHILOSOPHY

Vol. 11, Issue 2, 2019

The issue is sponsored by the Bulgarian National Science Fund, competition "Bulgarian Scientific Periodicals – 2018", project No КП-06-НП/55 from 12 Dec. 2018

TABLE OF CONTENTS

Articles

Discussion

Book reviews

BALKAN JOURNAL OF PHILOSOPHY. Vol. 11, Issue 2, 2019

Articles

CONSCIOUSNESS: THE POINT OF VIEW OF PROCESS PHILOSOPHY AND GENETIC STRUCTURALISM – A CRITICAL COMPARISON AND SOME CONSEQUENCES*

Franz Riffert

Department of Educational Science, University Salzburg,
Erzabt Klotz Straße 1, 5020 Salzburg, Austria,
email: Franz.Riffert@sbg.ac.at

Abstract:
First a sketch of the current state of the debate of the phenomenon of consciousness is provided; based on it David Chalmer's distinction between the weak and the hard problem of consciousness will be introduced. It will be indicated that Whitehead's process philosophy is able to offer a promising basis for solving the hard problem by showing how the concept of consciousness is anchored in his metaphysical theory. In the remaining parts of the paper the so-called weak problem of consciousness will be addressed in more detail by comparing Whitehead's speculative account with Piaget's empirical research results concerning the phenomenon of consciousness. By showing that Whitehead's and Piaget's positions on consciousness overlap widely a kind of indirect empirical support of Whitehead's philosophical position will be achieved. At those points where the two positions deviate – that is in only two points – there are indications that Whitehead holds the more plausible position. So this investigation confirms William Seager's suggestion that more attention should be drawn to Whitehead's process philosophy when trying to solve the hard, as well as the weak problem of consciousness. In order to substantiate this claim two topics are discussed from the point of view of Whitehead's position of consciousness: (a) learning and consciousness and (b) artificial consciousness.

Key words: Whitehead, Piaget, weak problem of consciousness, development of consciousness, learning, artificial consciousness.

1 Introductory problem exposition

'Consciousness' is a fundamental human phenomenon and therefore a central topic for philosophy, but also for psychology and the neurosciences; today it is even becoming a topic of Artificial Intelligence (AI) research (Aleksander, 2005 & 2007; Manzotti & Chella, 2013; Seth, 2009). Since Descartes' radical divorce of res extensa (material entities) and res cogitans (mental entities), it was also tightly connected with the so-called Mind-Body problem (Pothast, 1987; Velmans & Schneider, 2007).

In the first half of the twentieth century, due to the rising influence of Neopositivism and its psychological counterpart, radical Behaviorism, the concept of consciousness came under general suspicion of being unscientific and, as a consequence, vanished from empirically oriented scientific discourse, with the exception of psychoanalytic circles (which too were attacked as being pseudo-scientific).

The contents of consciousness are unavoidably subjective; what happens within a human person – when, for instance one tries to solve a chess problem – is only accessible to the person in question. This interior first-person perspective cannot – or if, only with great losses – be grasped from the third person perspective; therefore no objective, or at least intersubjective tests of these mental contents are possible. This fundamental situation makes the phenomenon of consciousness for empirical-experimental sciences, for instance, psychology and the neurosciences, very difficult to adequately examine. Consequently, 'consciousness' vanished as a topic of research for several decades during the twentieth century. Oeser and Seitelberger (1988, p.23, my translation) state for instance: "After the publication of

* This article is supported by the Bulgarian Science Fund according tot he project No DH 15/9 from Dec.2017.

Ryle's book [The Concept of Mind, 1943] there was no theory of mind for more than a decade."

Even worse was the situation for the complement of consciousness, the notion of the so-called 'unconscious'. During the 1950s the use of this word was frowned upon by those who then considered themselves true scientists. The mere speculation that a causal relationship between observable events and unobserved subliminal stimuli might exist was considered scandalous (Smith, 2004, p.27). Indeed, it remained woefully unclear how a researcher might apply the analytical principles of the scientific method toward the investigation of a concept that per definitionem could not be grasped consciously. On the other hand, it appeared impossible to elucidate the phenomenon of consciousness discursively while excluding the interrelated concept of the unconscious from scientific research.

During the last thirty to fourty years, the winds have changed dramatically: about a quarter of a century ago (1993) the first issue of the *Journal of Consciousness Studies* was released and has since served as a highly respected platform for discussing manifold contents related to consciousness (including all possible forms and levels of unconscious processes). After this breakthrough the number of scientific publications dealing with various facets of consciousness has exploded. "Today, we can see a steady rise in empirical articles that cite consciousness or closely related terms, currently 5.000 articles per year." (Baars, 2007, p. 236).

The same is true for book releases. In 2007, to pick out just one example, *The Blackwell Companion to Consciousness* was published (Velman & Schneider). The editors of this approximately 700-page anthology have aimed to provide a multidisciplinary overview of the current state of scientific research related to the burgeoning field of consciousness studies. Nearly seventy authors have contributed rich, concise articles on diverse topics, including the problem of defining consciousness as a subject of academic investigation. This problem comprises many subsidiary questions and areas of research. These include, for example: the problem of gradations of awareness (as opposed to a binary understanding of unconscious and conscious states of awareness in human subjects); the nature of the relationship between conscious and unconscious mental processes; the biological origins and neural correlates of conscious and unconscious processes; the variety of manifestations of conscious and unconscious subjective experience (ranging from self-reported meditative and mystical experiences to the temporary suspension of conscious awareness during medically-induced anaesthesia and the potentially permanent loss of conscious awareness in comatose individuals); pathological manifestations of conscious experience, such as 'blind sight,' retrograde and anterograde amnesia, and altered mental functioning/neurological abnormalities (or "split-brain syndrome") in individuals who have experienced partial or complete severing of the Corpus Callosum, which connects the right and left hemispheres of the brain; the so-called "binding" problem between mind and brain; the role of attention in generating conscious perception; and the role of neuronal correlates–the minimal set of neuronal mechanisms and events necessary for bringing about and sustaining conscious cognitive processes. Contributors to the anthology also examine the role that quantum physics might play in explaining consciousness, address the now classic Mind-Body Problem, and explore the results and implications of Benjamin Libet's neuroscientific experiments on the nature of free will.

Thus, twenty-five years after David Chalmers declared, "consciousness is making a comeback" (1994, p.22), his prediction has been fully confirmed. This trend was fueled by some early, but important research results. One of the most important catalysts for the renewed interest in the topic of consciousness in theoretical and empirical research was Piaget's genetic epistemology (1937, 1973). George A. Kelley's Psychology of Personal Constructs (1955), and Miller, Gallagher, and Pribram's Plans and the Structure of Behavior (1960) have also been influential in reviving an academic interest in consciousness. The aforementioned researchers have all undermined radical behaviorist positions by demonstrating (1) that there exist reliable methods of scientifically investigating subjective (first-person) experience and, (2) that this knowledge contributes appreciably toward our understanding of human beings. At about the same time in the philosophical domain, Karl Popper (1966 1974) – in collaboration with John Eccles (1977) – paved the way in the same direction; and, of course, Thomas Nagel's highly influential article "What it is Like to be a Bat" (1974) need to be mentioned in this context, as well as Mario Bunge's *Body-Mind Problem. A Psycho-Biological Approach* (1980).

Neuroscientists also testified to the importance of investigating the nature of consciousness when they discovered strong evidence indicating that un-

conscious processes may influence human action. Two key examples of this research include: (1) "Split-Brain experiments" (Sperry, 1971; Gazzaniga 1970) and, (2) the pathological "blind-sight" phenomenon (Weiskrantz 1986, 1997).

In experimental psychology early support came from micro- (the University of Leipzig'sGanzheits-psychology: Werner, 1956; Riffert, Bröderbauer & Huemer, 2015) and percept genetic approaches (Lund University: Kragh, 1955; Smith, 1957). One of the heatedly discussed papers was Antony Marcel's "Conscious and Unconscious Perception: Experiments on Visual Masking and Word Recognition," (1983) which challenged traditional theories of reading and convincingly demonstrated – by using the so-called forced choice method – the impact of subliminal stimuli on word recognition.

Broad empirical support for the existence of conscious and unconscious phenomena has also triggered research on the distinctive characteristics of these two channels of information processing. While unconscious processes are pre-attentive, passive-automatic, and pre-symbolic modes of information processing, conscious processes are characterized by intentionality, attention, and the use of symbols (Dienes & Perner, 1999; Lieberman, 2003; Deutsch & Strack, 2006; Evans, 2009). Scientists also addressed the urgent question of how these two ways of accessing the world are connected to each other, if at all. The majority of approaches, which aim at answering this question, can be characterized as hybrid cognitive architectures (also termed two process or two system theories). Such hybrid cognitive architectures have become quite popular in many domains of cognitive science, including: memory (Schacter, 1987; Schacter & Chiu 1993; Smith & DeCoster ,2000, Barrett, Tugade & Engel, 2004); perception (Marcel, 1983; Epstein & Pacini, 1999; Normann, 2002); problem solving and decision making (Sloman ,1996; Sun, Slusarz & Terry 2005; Toates, 2006; Dijksterhuis & Aarts, 2010; Kahneman, 2011; Evans, 2012), artificial grammar learning (Reber, 1967, 1993), social psychology (Devine, 1989; Chaiken & Trope, 1999; Smith & Collins, 2009) and second language acquisition (Krashen, 1981; Schmidt, 1994; Ellis et al., 2009), as well as in neuroscience (Berry, 1993; Damasio, 1994; Fu, Q., Bin, Dienes, Fu, X, & Gao, 2013; LeDoux, 1992; Lieberman, 2007; Milner & Goodale, 1995, 2008; Reverberi et al., 2005). Researchers in the domain of Artificial Intelligence have also adopted these hybrid cognitive architectures, using them for modeling human cognition (Anderson, 1983; Kokinov, 1997; Kokinov & Petrov, 2001; Sun, 2016). Ron Sun has even used his model as a basis for investigating the possibility of developing artificial consciousness (Sun, 1999, 2001).

In short, consciousness, and its counterpart the unconscious, have not only returned to the scientific sphere, but have come to play increasingly important roles within it. However, despite the resurrection of the topic of consciousness, many of the problems associated with it have not yet been resolved, and some of them remain far from any solution.

David Chalmers (1995a) has introduced a distinction between the conceptual and the empirical aspect of the problems of consciousness, allowing for a crude initial orientation. According to Chalmers, while the empirical problems associated with the concept of consciousness may prove more easily soluble, the conceptual problems present a greater challenge. Consistent with this distinction, Chalmers believed empirical problems, including those regarding the neural correlates of conscious phenomena, could be solved by applying the standard methods of the cognitive and neural sciences. The situation concerning the hard problem of consciousness, however, is quite different according to Chalmers. Indeed, the conceptual problems it raises have thus far successfully resisted any attempt to solve them within the classical physicalist (mechanist) paradigm–the metaphysical theory (often implicity) presupposed by the sciences.

This discouraging situation has increasingly led to calls for more radical ways of addressing these problems, including an overhaul of the established framework for investigating them, or a paradigm shift (Seager, 2003). The core of the hard problem consists in explaining how inner, subjective, and partially-active experiences and feelings, which at least sometimes reach the level of consciousness, can develop from a purely insentient, physical, and biochemical basis. Chalmers came straight to the point: "Remarkably, subjective experience seems to emerge from a physical process. But we have no idea how or why this is" (1995b, p.81). In light of persistent failure to solve the hard problem within the standard physicalist paradigm, Thomas Nagel cautiously stated: "Physical science has progressed by leaving the mind out of what it tries to explain, but there may be more to the world than can be understood by physical science" (1987, 36f).

Emboldened by doubts in the traditional paradigm, calls for radical paradigmatic reorientation in

the domain of consciousness studies have gained in vehemence during the last three decades. Yet, despite this situation, we still lack detailed theoretical approaches to challenge the status quo. Attempts offered so far are, at best, well-inspired sketches, often stuck at the level of metaphorical description. Chalmers, although aware of this situation, was not able to provide anything beyond crude speculation, stating, for example: "The full details of such a theory [challenging the accepted paradigm] are still out of reach, but careful reasoning and some educated inferences can reveal something of its general nature. [….] I propose that conscious experience be considered a fundamental feature, irreducible to anything more basic" (1995a, p.80, p.83).

About ten years later, Galen Strawson's seminal paper, "Realistic Monism: Why Physicalism entails Panpsychism" (2006), took Chalmers' intuition further and triggered an intensive discussion concerning such a panpsychist paradigmatic turn. In a special issue of the journal *Consciousness Studies* (2006, 13 (10-11)) Strawson's article was discussed by seventeen leading researchers. Its core idea, which Strawson summarizes as follows, is remarkably close to the basic intuition of Whitehead's process philosophy, which remains one of the most elaborate metaphysical theories to examine the problem of consciousness: "All physical stuff is energy, in one form or another, and energy, I trow, is an experience-involving phenomenon. This sounded crazy to me for a long time but I am quite used to it now and I know there is no alternative short of 'substance dualism', a view for which (as Arnauld saw) there has never been any good argument" (Strawson, 2006, p.25f). David Chalmers' and Galen Strawson's proposed solutions to the hard problem come close to an integral feature of Whitehead's process philosophy, namely, that dynamic "drops of experience" serve as the fundamental building blocks of reality (Whitehead 1978, p.18). Whitehead notably refers to these building blocks as 'feelings,' highlighting their subjective, inner character, and thus, the subjective character of reality as such.

Reduced to its core, Whitehead's proposed solution to the hard problem of consciousness consists of two steps. First, he develops a new metaphysics (a new paradigm), which ascribes to all its fundamental units (i.e. to actual entities) subjective sentience, or 'subjective form' as he terms it. In a second step, Whitehead shows how the specific subjective form of consciousness emerges during the process of change, or becoming, generating increasingly complex actual entities. Although actual entities at each level of metaphysical complexity have subjective form, the conscious character of the subjective form of (certain) complex actual entities presupposes the development and use of symbols, and therefore of a highly elaborate system of inner complexity. Hence, only certain complex actual entities are characterized by consciousness. Whitehead's metaphysics should, accordingly, be called 'panexperientialist,' rather than 'panpsychist' (Griffin, 1998).

The conceptual reorganization manifest in Whitehead's process metaphysics opens surprising new possibilities for solving long-standing problems; This includes providing a potential solution to the hard problem of consciousness, which, at its core consists in demonstrating how consciousness can develop from insentient, unconscious, and inert matter. The process philosophic position provides a plausible ontological solution to this puzzle, postulating that subjectivity (not consciousness!) is itself omnipresent in reality. What remains to be explained is how one specific form of subjectivity, namely consciousness, develops. In answering this question,Whitehead offers detailed suggestions, which I will examine in the following comparison between Whitehead's and Piaget's accounts of consciousness and its development.

In brief , it should come as no surprise that William Seager, one of the leading researchers in the field of consciousness studies, advocates an intense exploration of Whitehead's process metaphysics. As Seager states: "I would suggest that in the face of the evident difficulty of the problem of consciousness and the apparent convergence of several lines of argument towards some of Whitehead's most fundamental ideas, that it is time to have another serious look at Whitehead's approach" (2003, p.344).

Following Seager's suggestion, I will examine the ways Whitehead's process philosophy deals with certain problems relating to consciousness that are subject to empirical testing. In particular, I will illustrate how, according to Whitehead, consciousness (i.e. the specific form of subjective feeling) develops, and subsequently consider the relationship between Whitehead's theoretical account and the empirical research results Jean Piaget presents on this topic. In doing so, I will shed light on some surprising parallels between Whitehead's philosophical position and Piaget's empirical research results, indicating that Whitehead's account of the development of consciousness is–at least indirectly–supported by empirical evidence.

2. Comparing Whitehead's and Piaget's Accounts of the Development of Consciousness

Before going into detail, I find it necessary to briefly explain why I have chosen to survey Whitehead's and Piaget's respective accounts, and why I do not consider this pair an arbitrary choice for comparison.

First, it is important to note that both Whitehead and Piaget are proponents of the same paradigm. Indeed, Ludwig von Bertalanffy (1970) draws attention to the interesting historical fact that at the beginning of the twentieth century the so-called "mechanist world view" (1970, p.15 my translation) came under attack in different scientific disciplines, including: biology, experimental embryology, psychology, sociology, and cultural anthropology. At the same time, in the domain of philosophy, the Neo-Kantians also criticized the mechanist paradigm. Bertalanffy claims that as a result of this criticism the contours of a new paradigm emerged, which he terms the "organismic approach" (1970, p.14 my translation). He summarizes the fundamental characteristics of the organismic approach (in contrast to the mechanist paradigm) in the following way:

Treatment of the living system as a whole in contrast to a purely analytic and summative methodology; dynamic perspective contrary to a static or machine theory of organism; the basic assumption that organisms are primarily active and not reactive systems (Bertalanffy, 1970, p.14 my translation).

These three fundamental aspects of the organismic paradigm – holism, initiative activity, and dynamism – are readily identifiable in prominent features of Piaget's and Whitehead's theories. It is, therefore, no surprise that Bertalanffy considers "particularly Whitehead (1925) in philosophical respect […] a precursor" to the organismic paradigm (Bertalanffy, 1970, p.14 my translation). Interestingly, Bertalanffy also lists Piaget among major proponents of the organismic approach: "Again, at about the same time [when Heinz Werner and Ludwig von Bertalanffy were developing the organismic paradigm] the Genevan psychologist Piaget started his investigations into the mental development of the child in a way which can also be subsumed under the concept of the 'organismic'" (1970, p.15, my translation). Thus, Bertalanffy postulates a fundamental correspondence between Piaget's and Whitehead's approaches. This postulated correspondence can also be confirmed by reference to some passages in Piaget's oeuvre. For instance, in one of his books. (1972) Piaget not only positively refers to the usefulness of Bertalanffy's concept of 'organicism', but even makes explicit reference to Whitehead: "Already in his works on scientific thinking, Whitehead has taken the view that those interpretations that are commonly characterized as 'mechanic', cannot bring the analysis of reality to its utmost point, and that the terms 'organism' and 'organization' elucidate specific features that must be utilized" (Piaget, 1972, p.244, my translation).

Starting with these three fundamental consonances between Whitehead and Piaget, and motivated by the stimulating work of Reto Fetz (1984, 1999) I will advance several detailed comparisons between fundamental concepts of Piaget's genetic structuralism and Whitehead's process philosophy (Riffert, 1995, 2002). First, I will present Whitehead's reflections on consciousness. Second, I will introduce relevant findings from Piaget's empirical research that accord with Whitehead's position. Among several similarities that emerge from this study, the striking fact that the two scholars go so far as to formulate identical problems regarding the nature of consciousness will guide and, I hope, justify my project.

2.1 Whitehead on Consciousness

In the following section, I will investigate fundamental concepts of Whitehead's metaphysical theory that are critical to understanding his position on consciousness.

2.1.1 Whitehead's Fundamental Metaphysical Assumptions – a Sketch

Whitehead's organismic philosophy can be termed a one-type monism, since there is only one type of entity admitted in his theory, namely, actual entities, or actual occasions. Each actual entity is an occasion, or process, that creates itself by integrating, or incorporating the actual entities that immediately precede it through the prism of its own unique perspective. According to Whitehead, there is 'satisfaction' of this process of integration, or assimilation[1] (or, to introduce a neologism that Whitehead

[1] Here Whitehead clearly and explicitly departs from Aristotle's concept of first substances; whereas Aristotelian first substances cannot be predicated on other first substances, this is possible in Whitehead's metaphysical theory, since preceding actual entities are prehended (to introduce a Whiteheadian term) by the newly developing actual entity, and thereby integrated into a new actual entity. With the introduction of the concept of 'prehension,' Whitehead's approach comes very close to Piaget's

coins, 'concrescence') when it culminates in its final unity. At this point, the process of concrescence undergoes a change in ontological status, becoming a new actual entity that may function as an objective datum for other, inchoate subjective concrescences. Hence, the process is transformed into an entity that may be integrated into the matrix of a new perspective. In this way, a pulsating rhythm is generated that is constitutive of reality. Indeed, the rhythm of this pulsating net constructs the world as such as constant fluctuation between subjectivity (idealism) and objectivity (realism), and is, hence, the fundamental substratum of reality.

Although all actual entities are of the same type – i.e. have the same basic features – they can differ according to their complexity– i.e. according to their grade of inner differentiation. This is because a concressing actual entity can integrate the datum (the preceding actual entities) in different ways and to varying degrees of complexity. At one extreme there are the simplest actual entities that only passively reiterate the preceding actual entities and themselves show no active part in transforming their datum. At the other extreme lie human beings and their highly developed cognitive functions, which seem to be constituted by the most complex actual entities (that we have, as of yet, identified) in this cosmic epoch. In such high-grade actual entities, which form persons, the character of 'self-causation' is more strongly present than in the low-grade actual entities–that is to say, high-grade actual entities integrate and transform their datum more actively and in ways they tend to be more complex, dynamic, and innovative than low-grade actual entities.

Groups of actual entities may form so-called 'societies' of varying degrees of complexity. Such societies ultimately account for macroscopic entities, such as trees, dogs, and human brains. All actual entities that are members of a particular society exhibit a common feature, the so-called 'defining characteristic' of the society in question. The defining characteristic is transmitted to all newly developing actual entities that are integrated, providing the actual entity with its distinctive and unified identity. Again, it is crucial to note, there are many different types of societies. A chimpanzee, for instance, is a complex assemblage–partly hierarchically structured–of many fine-tuned societies. . The intricate systems that constitute cells, organs, bones, synapses, molecules, and atoms that make up the chimpanzee must be considered as societies of actual entities. Moreover, certain exceptional forms of societies even provide an environment whose conditions are favorable for the development of consciousness.

As far as we presently know, this may be the case for higher forms of animal life, including human beings, anthropoid apes, and some bird species.[2] Of course, not all actual entities that form complex societies generate conscious awareness. On the contrary, only occasionally and under advantageous conditions do non-human animals achieve higher levels of consciousness.

The development of consciousness presupposes the existence of what Whitehead calls 'personal societies,' which are only made possible due to the unique features of the animal brain. Personal societies are serially ordered societies–i. e. individual members of these societies succeed one another in sequential order (although, in the process, they may 'overlap,' or exist simultaneously). Members of such personal societies can, at times, reach the level of consciousness. Indeed, in humans, the integration of these interwoven and (partially) hierarchically structured personal societies constitute the 'human person,' including, of course, their body. Among other things, personal societies account for the uniform directedness (intentionality) of the individual's activities (goals, etc.). However, their influence over the complex ensemble of the human being is far from complete; sub-societies exist and may operate independently of the presiding personally ordered society (for example, the heart continues to beat when a person is in a state of coma, or even outside of the body, under specific artificial conditions). Personal societies also secure the diachronic identity of the individual human being by constantly transmitting and inheriting a defining characteristic.

Still, there remains a central task for human beings to contend with: in addition to establishing and maintaining diachronic identity, individuals must learn to adapt flexibly to their constantly changing surroundings, since "all societies require interplay with their environment" (Whitehead, 1929/1978, p.105). According to Whitehead, they reach this goal by developing "originality of response to stimuli" (Whitehead, 1929/1978, p.104). Whitehead recognizes that a simple, serially ordered personal society

genetic structuralism with its central concept of assimilation (Riffert, 1995).

[2] We may consider, for instance, the example of New Caledonian crows, which have been observed solving complex cognitive tasks and are capable of using and even creating tools.

will not be flexible enough to reach this second important goal since, in continuously reinscribing the (unchanging) defining characteristic, it permits itself to be constrained by the past. On the other hand, if flexibility is overemphasized, the endurance of diachronic identity is jeopardized and may break down, as demonstrated by well-known clinical cases of 'multiple personalities' (i.e. Dissociative Identity Disorder) and acute episodes of schizophrenia (Whitehead, 1929/1978, p.103). Hence, "[w]hat has to be explained is [how] originality of response to stimul[i]" is possible for human beings who simultaneously achieve (relatively stable) diachronic identities (Whitehead, 1929/1978, p.104). In order to do so, Whitehead introduces the concept of a so-called "*living* society" (Whitehead, 1929/1978, p.103 emphasis mine).

A personal society is called 'living' if it may interact freely with an "entirely living nexus" that provides the personal society with potential novel experiences and stimuli to which it may respond (Whitehead, 1929/1978, p.103).Since an "entirely living nexus is not a 'society'"–i. e. it is not able to sustain any defining characteristic by itself– it needs the protection of a suitable environment, which, in the case of human beings, is provided by the structured society of the brain (Whitehead, 1929/1978, p.102; pp.101-105). While Whitehead helpfully orients his reader with these particulars, he provides only a very rough sketch of his proposed explanation for the coexistence of the respective capacities for (1) retaining diachronic identity and (2) successfully adapting to environmental changes, through "origination of novelties of conceptual reaction" (Whitehead,1929/1978, p.102). These novel conceptual reactions allow the individual to incorporate new elements brought about by changes in the environment with "peculiar completeness" (Whitehead, 1929/1978, pp.101-105).–Instead of excluding disturbing environmental changes, the personal society integrates them by altering the pre-existing defining characteristic such that the old template (i.e. the defining characteristic) and new elements are united and form a new pattern.

But *where* is the *process of becoming conscious* to be located within the process of concrescence? And *how* is it achieved? Figure 1 provides a schematic presentation of three different major stages of integration of an actual entity (indicated in Figure 1 by their respective final culminations C, D, and E) which can take place in the process of concrescence.

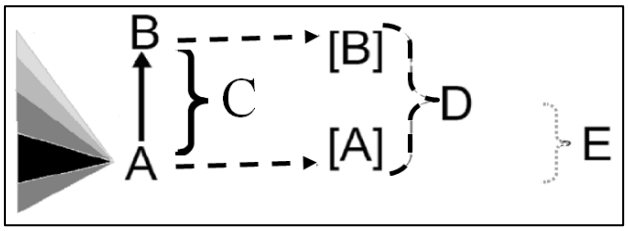


Figure 1. Schematic presentation of the different phases of a process of concrescence in a complex actual entity and the corresponding three modes of perception (For a similar scheme see: Sherburne, 1966, p.40.) First, the most primitive process of integration in a concrescence (i.e. the perceptive mode of causal efficacy) is described by the relation between A (physical prehension), B (conceptual prehension), and their reintegration in C (physical purpose). The relation between A, B, [A], [B] and D represents a so-called propositional prehension (perceptive mode of presentational immediacy). The most complex form of integration holds between A, B, [A], [B], D, and E (perceptive mode of symbolic reference). Consciousness can only arise in this latest, most complex form of concrescence, which Whitehead also terms intellectual prehension, and which is generally called 'judgment' in philosophy.

perspective of prehension A

A…perspective physical prehension(s) of immediately preceding actual entities (which are the datum of the concrescence) by physical prehensions

B…conceptual prehension, grasping an eternal object

C…physical purpose (synthesis between A and B; full, black, thick line)

[A]…logical subject (derived by reduction from physical prehension A)

[B]…predicative pattern (derived by reduction from the conceptual prehension B)

D…simple conformal prehension (propositional prehension; synthesis between [B] and [A]; dashed, gray, thin line)

E…intellectual prehension (synthesis between D and A; dotted, light gray, very thin line)

The first, most primitive level of integration that occurs in a process of concrescence starts (in its simplest version) with a physical prehension (A), which assimilates or prehends the immediately preceding actual entities from the particular perspective of this actual entity (which thereby initiates the process of self-constitution). From this physical prehension (A) a conceptual prehension (B) is derived – that is, the new actual entity that is coming into being grasps an eternal object that is present in the datum physically prehended by A. This primitive form of integration ends as soon as the physical prehension (A) and the conceptual prehension (B) are

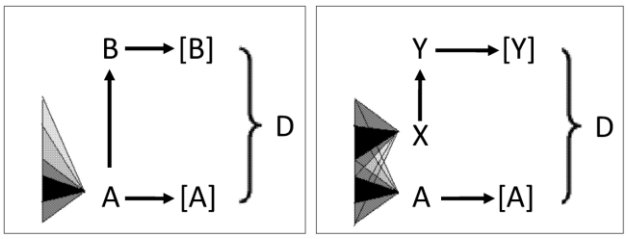
reunited in a so-called physical purpose (C) – in Whitehead's terms: it reaches its satisfaction and thereby turns into a partial datum for a newly developing actual entity. This primitive form of concrescence is the metaphysical correlate of the primitive mode of perception that Whitehead calls "causal efficacy" (1929/1978, p.176).[3]

If there is no re-unification between the physical prehension (A) and the conceptual prehension (B), instead a further differentiation and re-integration takes place, leading to what Whitehead terms "propositional prehension" (1929/1978, p.260). There are two different important propositional prehensions that are represented schematically in Figure 2, the perceptive propositional prehension: Figure 2 (a), which is identical with the perceptive mode of presentational immediacy, and Figure 2 (b), the so-called imaginative propositional prehension, which is isomorphic with the onset of human thought (Whitehead 1929/1978, p.261). This indicates that the multilevel process of integration serves as the metaphysical underpinning of both Whitehead's perceptive mode of presentational immediacy (Figure 2 (a)), and human thinking, generally (Figure 2 (b)).

Now, in propositional prehensions a "double elimination" or "reduction" takes place: (a) the physical prehension (A) will be reduced to a so-called "bare it" ([A]), which functions as the logical subject in a proposition, and (b) the conceptual prehension (B) will be reduced to a so-called "predicative pattern" ([B]), which is predicated of the logical subject (Whitehead 1929/1978, p.261, 276, 261, 261). Through this process a conceptualization takes place.[4]

While in a perceptive propositional prehension the eternal object (B) grasped by a conceptual prehension is derived from the physical prehension A (see Figure 2 (a)), in the case of an imaginative prehension the eternal object (Y) of an imaginative prehension is derived from a different physical prehension X (see Figure 2 (b)).

Here it is important to note that, according to Whitehead, the conceptualization–which occurs when the integration within the process of concrescence produces a propositional prehension–must be distinguished from a judgment. A judgment involves the elements A, B, [A], [B], D, *and* E, presented in Figure 1 (dotted, light gray, very thin line): in addition to the integration process of a propositional prehension, one further step of re-integration is added, and the propositional prehension D is re-integrated with the original prehension A, thereby achieving E (a judgment). Only actual entities that reach this high grade of internal differentiation and (re-)integration are capable of becoming conscious, either in (1) perception (through the re-integration of a perceptive propositional prehension with the original physical prehension) or (2) conscious thought (through the re-integration of an imaginative propositional prehension with the original physical prehension).



a) b)

Figure 2.
a) Perceptive prehension.
perspective of prehension A
A…perspective grasp of immediately preceding actual entities (which are the datum of the concrescence) by physical prehensions
B…conceptual prehension
[A]…logical subject (derived by reduction from physical prehension A)
[B]…predicative pattern (derived by reduction from the conceptual prehension B)
D…simple conformal prehension (propositional prehension)
b) Imaginative prehension
perspectives of prehensions A and X
A…perspective grasp of immediately preceding actual entities (which are the datum of the concrescence) by physical prehensions
X… perspective grasp of immediately preceding actual entities (which are the datum of the concrescence) by physical prehensions
Y…conceptual prehension
[A]…logical subject (derived by reduction from physical prehension A)
[Y]…predicative pattern (derived by reduction from the conceptual prehension B)
D…simple conformal prehension (propositional prehension)

[3] Riffert (2003) grapples with the question of whether or not Whitehead's bold postulation of the primitive perceptive mode of causal efficacy is supported by empirical evidence.

[4] Here we are at that point of transformation which computational scientists (for instance Smolensky (Smolensky & Legendre 2006) and Sun (1999, 2001)) are wrestling with: the passage from connectionist (distributed) processing to symbolic (localist) processing.

Consciousness reaches its peak when there is a contrast between propositional prehension (D) and physical prehension (A); i. e. when the conceptualization of the (particular section of the) world is false. That is, if our conceptions lead us to anticipate a certain (shape of) reality that the physical prehension belies, we become acutely attuned to, and therefore conscious of, this contrast. Whitehead (1929/1978, 24) refers to this phenomenon as the "'affirmation-neagtion' contrast". If there is no such contrasti.e. if our expectations of the world turn out to be correct,), we normally navigate through reality more or less smoothly, without many distortions, and hence, without much consciousness.

This short introduction into Whitehead's metaphysics should provide a basis for understanding the following theses, which will explain his theory of the development of consciousness in greater detail.

2.1.2 Six Thesis on Consciousness – the Whiteheadian Point of View

In what follows, Whitehead's position on consciousness will be presented in six theses.

Thesis 1: Primacy of experience over consciousness

Whitehead is very clear that "consciousness presupposes experience and not experience consciousness" (1978, p.53). According to Whitehead, humans find themselves involved in interactions with their environment in a variety of ways, which presupposes that they must in some way perceive their environment. The environment provides many positive challenges and opportunities, but also frightening threats, which humans often become aware of only after they are confronted with them, and sometimes only after they have started to react to them. After consciously tackling these challenges and threats, individuals often sink back to a subconscious level of interaction with their environment. Thus, most of our processes of engagement with the world take place at a subliminal level, before we become conscious of them. As Whitehead states:

> Our consciousness does not initiate our modes of functioning. We awake to find ourselves engaged in process, immersed in satisfaction and dissatisfaction, and actively modifying, either by intensification, or by attenuation, or by the introduction of novel purposes. *This primary procedure which is presupposed in consciousness*, I will term Instinct. It is the mode of experience directly arising out of the urge of inheritance, individual and environmental (Whitehead 1933/1967, 46, emphasis mine).

Consider, as an example of this, the so-called 'cocktail party effect,' which many of us have experienced at one time or another (Cherry, 1953; Broadbent, 1954). One is attending a party and absorbed in conversation with a small group. Suddenly, one hears one's name mentioned in a nearby group, despite having heard nothing of this group's conversation before. One's (singularly meaning-laden) name is the only word that has reached the supraliminal level of consciousness, and this has occurred even though the individual has not consciously followed the other group's conversation.

It remains outside the scope of my discussion to investigate why the individual's name has this privileged status–as the only word to reach the level of conscious awareness in this scenario–although this presents a fascinating topic of research. The cocktail party effect is, however, relevant to my project as it exemplifies the phenomenon of subliminal experience–in this case, auditory perceptions–preceding consciousness. If we were not continually immersed in subconscious interactions (including phenomenological experience) with our environment, it would be difficult to explain why we at times become conscious of certain stimuli, in particular, those that carry special meaning for us. Hence, given the ubiquity of subliminal experience, Whitehead maintains, "consciousness is the crown of experience, only occasionally attained, not its necessary base" (1978, p.267).

Whitehead's position, of course, has an impact on philosophy in general and on metaphysics in particular. Among other things, Whitehead's process philosophy indicates that philosophers who choose conscious human knowledge contents, which are given '*clare et distincte*' (Descartes), as the starting point of their investigation of human beings, have put the cart before the horse, so to speak, since conscious knowledge offers access to only a very limited slice of the full range of human experiences. Since such slices are strongly biased toward abstracted features, this decision strongly selects the basis of the philosophers' theories. Consciousness is a late comer in biological evolution, and much empirical evidence suggests that its development was favored because it increased the fitness of a species for coping with very specific everyday tasks that were critical for survival. So, if one is dealing with topics beyond everyday issues of survival, as is often true in the case of metaphysics, an uncritical trust in

consciousness as the natural starting point of one's philosophical theory may easily lead to error and confusion.

Conscious clarity, according to Whitehead, is *the aim* of philosophy (particularly metaphysics), and of the sciences, but is should not and cannot be the starting point of their investigations (Riffert, 2012). Indeed, as Whitehead states,

> [c]onsciousness primarily illuminates the higher phase in which it arises, and only illuminates earlier phases derivatively, as they remain components in the higher phase. Thus those elements of our experience which stand out clearly and distinctly in our consciousness are not its basic facts; […] For example, consciousness only dimly illuminates the prehensions in the mode of causal efficacy [instinct], because these prehensions are primitive elements in our experience. But prehensions in the mode of presentational immediacy are among those prehensions which we enjoy with the most vivid consciousness. […] The consequences of the neglect of this law, that the late derivative elements are more clearly illuminated by consciousness than the primitive elements, have been fatal to the proper analysis of an experient occasion. In fact, most of the difficulties of philosophy are produced by it. [...] It follows that the order of dawning, clearly and distinctly, in consciousness is not the order of metaphysical priority (1978, p.162.)

Thesis 2: Consciousness does *not* consist in a mere process of illuminating, which would leave unchanged that which is illuminated

Prima facie, the casual reader could easily develop the mistaken impression that Whitehead believes 'becoming conscious' consists in the mere illumination of a content already present to the individual at a subconscious level, and that a content previously registered on a subconscious level remains unaltered when the individual comes to grasp it consciously. This, however, is not Whitehead's position.

For Whitehead, consciousness is the subjective form of a new type of prehension, the so-called 'intellectual prehension.' Intellectual prehensions (re-)integrate more primitive forms of prehensions, rather than simply repeating them. The contents of intellectual prehensions are thus different from the contents of simpler, more primitive prehensions. To introduce a term from Piaget's work: the process of becoming conscious of some content (that we have hitherto registered at only a subconscious level) necessarily involves its 're-construction.' To put it metaphorically: becoming conscious is not comparable to switching on stage lights and making visible to the audience sets that were always already there. On the contrary, only by integrating – and therefore changing – more primitive elements do they become 'illuminated,' or consciously available to us. To restate Thesis 2 in positive terms: whenever a previously unconscious datum is consciously grasped, it is transformed by the very process of being made available to consciousness.

Thesis 3: Consciousness presupposes creation of propositions (i. e. symbols).

According to Whitehead, propositional prehensions can be divided into (a) perceptive propositional prehensions[5] and (b) imaginative prehensions. Intellectual prehensions presuppose propositions (and propositional prehensions) that are held in contrast to physical prehensions. Intellectual prehensions (α) can therefore be divided into (α1) conscious perceptions, which presuppose perceptive propositional prehensions (a) as elements, and (α2) intuitive judgments, which integrate imaginative propositional prehensions (b).

Further, propositions symbolize physical prehensions (which are not necessarily simple), and may translate these verbally (although this is not necessarily entailed). For the sake of clarity, verbalized propositions offer obvious advantages over other symbolic translations of physical prehensions. The grammatical subject of a verbalized statement stands for the logical subject of a proposition, while the grammatical predicate represents the so-called "predicative pattern" of the proposition in question (Whitehead, 1978, p.257). However, we must bear in mind that the *verbal* expression in its entirety–the sentence–stands for the proposition, while the symbolic proposition stands for the fact of experience that it communicates. Hence, a verbalized proposition only indirectly symbolizes a specific section of reality (if, of course, the proposition in question has a positive truth-value).

A proposition is a symbol that refers to a physical prehension, which itself has immediately grasped

[5] However, this is only a crude differentiation, since perceptive prehensions can be divided further into (a_1) authentic (veridical) and (a_2) unauthentic (illusory) perceptive prehensions.(Whitehead, 1929/1978, pp.262-270; see also Mays, 1977).

reality from its own unique perspective. In an intellectual prehension, this relation between propositional prehension and (the more primitive) physical prehension is explicitly realized, creating a contrast between propositional prehension (reality as it might be) and physical prehension (reality as it in fact is).The proposition in a propositional prehension thus symbolizes a hitherto unconscious primitive physical prehension, and it is the integration of these two elements (i.e. the physical prehension and the symbolic proposition) that allows the fact of experience to rise the level of conscious awareness. Consequently, "there is no consciousness apart from propositions as one element in the objective datum" \of an intellectual prehension (Whitehead, 1978, p.243).

Thesis 4: Consciousness is created by the prehension of the 'affirmation-negation' contrast.

By now it should be clear that, according to Whitehead, consciousness presupposes propositions (and propositional prehensions), which symbolize physical prehensions. While symbolization is *necessary*, it is not a sufficient condition for the creation of consciousness. According to Whitehead, propositional prehensions are not conscious; consciousness comes into play *only* in a judgment (intellectual prehension), which affirmatively integrates the propositional prehension with a primitive physical prehension, thereby affirming the proposition's positive truth-value. Thus, the primitive physical prehension functions as (an aspect of) 'reality'–that which is the case–and the proposition functions as a sign that may (or may not) accurately symbolize, or signal, this (aspect of) reality. In this respect, an intellectual prehension creates "the contrast between '*in fact*' and '*might be*', in respect to particular instances in *this* world. The subjective form of the prehension of this contrast is consciousness" (Whitehead. 1978, p.267). This contrast is fully manifest if the symbol (which refers to what might be) does not correctly (veritably) symbolize the fact (of what is).

> Whitehead presents the following example to illustrate these ideas: a person verbally utters the proposition, "This stone is grey." If the person now approaches the stone from a different position it may turn out that the stone, in fact, is not grey (but instead, dark blue). In this moment, in which the affirming proposition is negated by the fact (physical prehension), the contrast between 'might be' (propositional prehension) and 'in fact' (physical prehension) reaches its full expression. Whitehead refers to this phenomenon as the "affirmation-negation contrast," and explains, The general case of conscious perception is the negative perception, namely,'perceiving this stone as not grey'. The 'grey' then has ingression in its full character of a conceptual novelty, illustrating an alternative. In the positive case, 'perceiving this stone as grey', the grey has ingression in its character of a possible novelty, but in fact by its conformity emphasizing the dative grey, blindly felt. Consciousness is the feeling of negation: in the perception of 'the stone as grey' such feeling is in barest germ; in the perception of 'the stone as not grey' such feeling is in full development. Thus the negative perception is the triumph of consciousness" (1978, p.243, p.161).

The same applies to (scientific) theories, since Whitehead conceives of them as complex propositions (Whitehead, 1978, p.184, p.186, and more often). While a singular proposition expresses a potentiality for a specific set of actual entities (or a specific type of nexus of actual entities), general propositions differ from them only "by the generalization of 'one definite set of actual entities' into 'any set belonging to a certain sort of sets'" (ibid, p.186). A general proposition, moreover, is termed 'universal', if "the sort of sets includes all sets with potentiality for that nexus of reactions" (ibid). Given this parallelism between singular propositions and theories (general propositions), it is no surprise that Whitehead can also claim for theories: "Consciousness requires more than the entertainment of theory. It is feeling of the contrast of theory, as *mere* theory, with fact, as *mere* fact" (ibid, p.188).

At this point it is of some interest to consider the condition Whitehead has identified as "suspended judgment" (1978, p.272). Whitehead maintains suspended judgments engender the possibility of running through many potentialities "which are not datively exemplified" in reality (Whitehead, 1929/1978, p.161). But how can these 'thought games', which do not aim at a particular (part of) reality, and which, therefore, seem to exclude the possibility of prehending a contrast, become conscious?

Whitehead solves this problem by referring to the possibility of a contrast between a theory and a potential reality, i. e. another proposition. Such a contrast between two propositions, which may end in a contradiction, may also lead to the creation of

consciousness. As Whitehead states, "It [imaginative, symbolic thought] can even play with inconsistency; and can thus throw light on the consistent, and persistent, elements in experience by comparison with what *in imagination is inconsistent* with them" (Whitehead, 1978, p.5, emphasis mine. As an example of the importance of such a 'play with inconsistencies,' we may consider the role of indirect proof in mathematics or logic. One 'playfully' assumes the opposite of something one holds as true and seeks to prove in order to demonstrate that this 'playfully' adopted assumption leads to contradiction, thereby proving (via the Law of Noncontradiction and the Law of Excluded Middle within a binary system of truth-values) that the opposite claim must be true. Piaget also applies this procedure in his research when he first (hypothetically) assumes that a one-year-old child should have developed the concept of 'object permanence,' only to then empirically demonstrate that, if this were, in fact, true, the child would exhibit very different patterns of behavior than those observed. This allows Piaget to claim with some confidence that one-year-old children have not yet developed the concept of object permanence.

Thesis 5: Consciousness is not an all-or-nothing phenomenon, but comes in gradations

Whitehead maintains there exist different states of consciousness, and therefore, consciousness as such is subject to gradation. In the example of the proposition, 'This stone is grey,' Whitehead hints at this notion, and affirms it explicitly in the following passage:

> We sleep; we are half-awake; we are aware of our perceptions, but are devoid of generalities in thought; we are vividly absorbed within a small region of abstract thought while oblivious to the world around; we are attending to our emotions – some torrent of passion – to them and to nothing else; we are morbidly discursive in the width of our attention; and finally we sink back into temporary obliviousness, sleeping or stunned (Whitehead, 1978, p.161).

According to Whitehead, the mere correspondence between the predicate pattern of a proposition and the fact to which it intends to refer triggers consciousness only "in barest germ" (ibid). However, when a proposition fails to represent a fact veridically, full consciousness is reached. Thus, "[t]he negative judgment is the peak of mentality" (ibid, p.5).

Thesis 6: It is the main function of consciousness to deepen aesthetic experience

The main function of conscious intellectual prehensions is "to heighten the emotional intensity" of phenomenological experience (Whitehead, 1978, p.272). Indeed, intensifying emotion is the major aim of *all* processes within Whitehead's metaphysical theory. Suspended intellectual judgments in particular open the possibility of playing with a wide range of combinations of airy speculations not realized in reality, thereby arousing hitherto unknown, exciting emotional experiences.

It appears Whitehead is skeptical of the oft-cited claim that consciousness has conferred adaptive evolutionary advantages on homo sapiens. According to Whitehead, it is unclear what effect higher-order thought processes may have on the survival of the human species; though complex new ideas may bestow evolutionary advantages, they are equally likely to invite disaster:

> The nearest analogy [for the power of conscious imaginative thought] is to be seen in the history of some species of animal, or plant, or microbe, which lurks for ages as an obscure by-product of nature in some lonely jungle, or morass, or island. Then by some trick of circumstance it escapes into the outer world and transforms a civilization, or destroys an empire or the forest of a continent. Such is the potential power of the ideas which live in the various systems of philosophy (Whitehead, 1933/1967, p.146).

Hence, while new ideas may be the products of finely calibrated thought processes that are unique to highly evolved animal species, this alone does not ensure they will prove advantageous. Rather,"[t]he novelty may promote or destroy order; it may be good or bad" (Whitehead, 1978, p.187).

It is by no means clear that the development of higher-order (conscious) thinking has increased human beings' fitness for survival. As Whitehead reminds us,

> So far as survival value is concerned, a piece of rock, with its past history of some eight hundred million years, far outstrips the short span attained by any nation. *The emergence of life is better conceived as a bid for freedom on the part of organisms, a bid for a certain independence of individuality with self-interests and activities not to be construed purely in terms of environmental obligations. The im-*

> *mediate effect of this emergence of sensitive individuality has been to reduce the term of life for societies from hundreds of millions of years to hundreds of years, or even to scores of years. The emergence of living beings cannot be ascribed to the superior survival value either of the individuals, or of their societies* (Whitehead, 1927/1985, pp.64-65, emphasis mine; see also 1929/1958, pp.4-34).

On the contrary, "[t]he problem set by the doctrine of evolution is to explain how complex organisms with such deficient survival power ever evolved" (Whitehead, 1929/1958, p.5). What, then, is Whitehead's explanation for this remarkable situation?

The solution Whitehead offers to this puzzle is utterly original and, by some accounts, heretical. Whitehead asserts there "is a three-fold urge: (i) to live, (ii) to live well, (iii) to live better. In fact the art of life is *first* to be alive, *secondly* to be alive in a satisfactory way, and *thirdly* to acquire an increase in satisfaction" (ibid, p.8).

Further details of Whitehead's thought on evolution exceed the scope of this paper. Based on the ideas discussed above, I hope to leave my reader with the following crucial takeaway: Whitehead (in contrast to Piaget, as we shall see)challenges the belief that the development of ever more complex thought patterns increases human beings' fitness for survival, maintaining instead that the chief result of higher-order, conscious thought processes is the arousal and intensification of emotional experience..

2.2 Piaget on Consciousness

In most of his books, Piaget addresses the topic of consciousness (for example: 1918, 1924/1969, 1936, 1937, 1950), and in a few publications he focuses on it exclusively (for example:*Reussir et comprendre* (1974a) [*Success and Understanding* (1978)], *Recherches sur la contradiction (*1974b) [*Experiments in Contradiction* (1980)], *La prise de la conscience* (1974c) [*The Grasp of Consciousness* (1976)], and *La formation du symbole chez l'enfant* (1959) [*Play, Dreams and Imitation in Childhood* (1962), *The Problem of Consciousness in Child Psychology: Developmental Changes in Awareness* (1954)]).[6] In each of these publications, Piaget describes his empirical investigations on the development of consciousness in children and draws theoretical conclusions based on these results.

Though Piaget is well-known for his research on cognitive development, his investigations on the topic of consciousness are still widely neglected, and "[e]ven today Piaget's conception of consciousness is not well known" (Pons & Harris, 2001,p. 220). Indeed, more than ten years after this initial charge, the same authors lament that "[a]mong the substantive works that synthesize Piaget's scholarship, his conception of consciousness is never fully presented and discussed" (Pons, Harris & Rosnay, 2012, p.96; for a recent counter-example see Stoltz, 2018). This is so, despite the fact that the topic of consciousness has returned to the center of lively scientific and philosophical debate. Given the strong impact of Piaget's work on the development of cognitive science, one may wonder if his research on consciousness could have a similar effect on consciousness studies.

In what follows the results of Piaget's research on consciousness. will be presented and discussed. However, only those aspects of Piaget's work will be presented which are relevant for comparison with Whitehead's position on the topic of consciousness, and which, therefore, may confirm or disconfirm Whitehead's view.[7]

Thesis 1: LikeWhitehead, Piaget claims that consciousness presupposes sub- or unconscious knowledge

From the moment of conception onward–i. e.during the prenatal phase up until the second year of postnatal development–human beings are involved in pre-conscious interactions with their immediate environment. At about age two, children begin to engage in certain specific interactions in a conscious manner; however, even then the broad majority of interactions remain at the subconscious level. Accordingly, Piaget distinguishes between two different modes of knowing: savoir-faire [know-how], or the non-conceptual knowledge or skills that allow one to act, and savoir, or conceptual knowledge. Savoir-faire, the primitive and unconscious mode of knowing, is the only mode of knowing available to the child prior to the development of conscious (intentional) interactions with his environment. Hence, in the beginning, all of the child's interactions with his environment occur at a subconscious level. Piaget's research on child development

[6] All quotations taken from these publications will be taken from the English translations.

[7] It is important to note that this paper neither aims at, nor claims to provide a complete exposition of Piaget's position on the phenomenon of consciousness.

shows "that in all the elementary successes, conceptualization lags behind action – clear proof that the latter is autonomous" (Piaget, 1974a/1978, p.214). This indicates that a conscious grasp of reality is founded on the more primary (and unconscious) coupling of actor and environment that occurs through savoir-faire, and is built up through an active process of reflective abstraction.

Piaget's distinction between savoir and savoir-faire helps us make sense of the fact that children can possess accurate unconscious knowledge, while lacking correct conscious knowledge. If, for instance, a group of eight-year-old children are asked to explain verbally (conceptually) how they crawl on all fours, the majority will answer that they move their arms and legs simultaneously, which, of course, is wrong. However, most will have no trouble crawling; i.e. most of these children possess the correct *savoir-faire* of crawling. Hence, according to Piaget, unconscious *savoir-faire* is temporally and logically prior to conscious-conceptual knowledge. On this point Piaget and Whitehead fully agree.

Thesis 2: The development of consciousness is an active process

Piaget continually draws attention to the fact that the process of becoming conscious is not a purely passive event of becoming aware of hitherto unconscious contents of *savoir-faire,* without actively altering these contents. Quite to the contrary: it is "clear that in no way can the process [of becoming conscious] be reduced simply to throwing light on these elements without, in doing so, modifying them" (Piaget, 1974c/1976, p.337).

For Piaget, as for Whitehead, the process of becoming conscious of previously unconscious mental contents entails actively transforming these contents through reflective abstraction.. This, to be precise, involves '(re-)constructing' unconscious contents at a higher level. As Piaget states, "If cognizance were merely the result of throwing light on a situation, these coordinations would necessitate no new construction, since they have already been achieved at the level of the material action itself (know-how as opposed to knowing)" (Piaget, 1974c/1976, pp.338-339).

The new quality to which the process of consciousness leads becomes particularly evident in those cases, in which humans are able to reconstruct the unconscious contents of *savoir-faire* only in an incorrect way. This can be illustrated by reference to Piaget's experiment concerning the path of an object launched by a sling (1974c/1976, 12-45): Children at the age of seven or eight are, after some training, able to successfully throw a wooden ball (with a diameter of 5 centimeters), which is attached to a sling, into a target (basket) positioned directly in front of them, by performing horizontal circular movements in a clockwise direction at ground level.The target is located at the position 12 o'clock, while the children are located at 6 o'clock (with regard to the circle described by the sling). The task can only be solved successfully if the sling is released by the child at a position of approximately10 o'clock (or at 2 o'clock, depending on the direction the child turns the sling).

Although children can learn to accomplish this task, if they are asked to describe what they are doing, and particularly, if they are asked to name the precise point of the circular movement at which they release the sling, they are not able to give the correct answer. In fact, most children at age seven or eight claim they release the sling at the 12 o'clock position, i. e. directly opposite of the target. This experiment demonstrates clearly that unconscious *savoir-faire* and conscious savoir are not congruent; they differ substantially, and conscious knowledge lags behind unconscious know-how. Becoming conscious thus presupposes active reconstruction processes, and by no means can be explained by a simple, passive illumination of unconscious mental contents that are pre-given. Here again Piaget and Whitehead are in agreement.

Thesis 3: The reconstruction of unconscious sensori-motor schemes takes place at the level of concepts (symbols)

Piaget asserts that, "[c]ognizance must, from the start, involve conceptualization" (1974c/1976, p.338). Through conceptualization, pre-conscious sensori-motor schemes are elevated to another (higher) level and are reconstructed there. As Piaget explains: "The subject looks at his[/her] actions and these are assimilated, more or less adequately, by his[her] consciousness as if they were ordinary material links situated in the objects – hence the necessity for new conceptual constructions to account for them. In fact only reconstruction is involved, but it is as laborious as if it corresponded to nothing already known by the subject him[/her]self" (Piaget 1974c/1976, p.339). If reconstruction of pre-conscious elements were unnecessary, and consciousness "were simply a mirror," then" it would therefore be sufficient for consciousness, if this were simply a mirror, to reflect objectively the hitherto unconscious movements of the action in order to

obtain a 'representation' (in the most direct meaning of the term) of the coordinations already being effected" (Piaget 1974c/1976, p.339). Thus, concepts are symbolizations of unconscious sensori-motor activities or of those aspects of reality that are assimilated to these schemes. For Piaget, as for Whitehead, the act of becoming conscious thus presupposes the so-called 'symbol function'. Since conceptual, or symbolic thinking develops around the age of two, it would appear that this, too, is the age when consciousness must arise.

While Piaget and Whitehead both identify a close connection between the development of consciousness and the development of conceptual thinking, Piaget posits an even stronger thesis, claiming that the respective processes of becoming conscious and of conceptualizing coincide– i. e. that they are identical, or are, in fact, one and the same process. Becoming conscious of something is the conceptualization of it, and the conceptualization of something is the act of becoming conscious of it. Becoming conscious "consists, from the outset, in a true conceptualization, or, in other words, in a passage from practical assimilation […] to an assimilation through concepts" (Piaget, 1974c/1976, p.337).

On this point, Whitehead and Piaget hold different positions. It seems that for Piaget–just as for Aristotle[8]–propositions (which conceptualize) and judgments coincide, and, indeed, Piaget explicitly states that the process of conceptualization is accompanied by consciousness (Piaget, 1974c, p.338).

Thesis 4: Becoming conscious is the result of an experienced contrast between a concept (symbol) and reality

According to Piaget, consciousness appears when a process of conceptualization takes place. Indeed, Piaget does not distinguish between these processes; the act of conceptualizing just is the act of becoming conscious. By contrast, according to Whitehead, concept formation is a necessary, although not a sufficient condition of consciousness arising; in addition to conceptualization, there must occur an act of judgment that establishes a correspondence between the concepts formed and the portion of reality they seek to capture.This is because, according to Whitehead, a subject can possess concepts *unconsciously, and* as long as there is no additional judgment (regarding the veridicality of these concepts), they will remain unconscious.

Piaget accepts Claparéde's 'law of becoming conscious' (1905), which holds that successful conceptualizations create no reason sufficient for the creation of consciousness, and additionally states: "[c]ognizance would thus appear a direct consequence of a failure to adapt and would be useless as soon as the functioning […] adapts itself normally" (Piaget, 1974c/1976, p.333). Following this logic, consciousness develops if and only if the process of adaptation to reality fails. Interestingly, Piaget illustrates his position in the book *Resussir et comprendre* (1974a/1978) by referring to an example that is surprisingly similar to the example Whitehead references–the proposition, 'the stone is grey'– when discussing the same problem: "[W]e perceive that an object is red, or square, or on top of another, etc. […] the negative qualities having no significance, except in relation to other objects used for comparison, to *predictions* that do not come true, or to momentary *needs* that are not satisfied" (Piaget, 1974 a/1979, p.230 emphasis mine). These contrasts between proposition (concept) and reality (or more precisely, the reality which is coupled with the subject's sensori-motor schemes) are decisive. Negations–Piaget continues–"involve correlations, coordinations and often increasingly complex inferences. We are therefore dealing with particularly important examples of the general progress of cognizance" (Piaget 1974a/1978, p.230f). Like Whitehead, Piaget remarks upon the contrast between concept and reality, and–more surprisingly still–even uses the same terminology of "affirmations and negations" to designate it (Piaget, 1974a/1978, p.228). These similarities compel one to ask whether Piaget was familiar with Whitehead's writings on consciousness. In either case, we can clearly identify many parallels between their respective positions.

Thesis 5: According to Piaget, concluding that consciousness is subject to gradation is inevitable

While Piaget's perspective on gradations of consciousness is consistent with Whitehead's position, Piaget approaches this topic from a different perspective.

During his empirical investigations, Piaget became aware that many of a child's early conceptualizations, or observations, of the world were subject to distortion and error. As an example of this phenomenon, let us return to the sling experiment. As we have seen, when children at the age of seven or eight are asked to throw a ball into a basket (target)

[8] Cataldo (1982) has drawn attention to this difference between Whitehead and Aristotle. He also has tried to show that a clear distinction between propositions (i. e. 'concepts' in Piaget's terminology) and judgments can be omitted, as can Aristotle's problem of 'negative facts.'

by manipulating a sling, they quickly learn to successfully complete this task. However, when asked to describe this activity, they maintain that they release the ball in a position directly opposite the target (a position of approximately 12 o'clock), despite the fact that they have released the ball at a position of approximately 10 o'clock in all of their successful trials. While they have learned to accomplish the task, their conceptualization of their own behavior is wrong.

This leads Piaget to question why the children are unable to correctly observe and conceptualize their activities. To make sense of this problem, Piaget first tests the hypothesis of cognitive "repression" (Piaget, 1974d, p.36; 1974c/1976, pp.340-342). According to this theory, the child has previously thrown a ball with his hands, and thus learned (consciously) that–under these conditions–he successfully gets the ball into the basket when he releases it from the 12 o'clock position. When asked to describe his successful approach to the sling problem in conceptual terms, the child falls into error by assuming that his consciously-held knowledge of accurately throwing a ball with his hands also applies to the sling problem–that is, he overgeneralizes this conceptual knowledge.. The conceptual knowledge the child has developed based on multiple occasions of throwing a ball with his hands appears well-corroborated by experience; yet, this patently accurate knowledge is the very thing that impedes the child's ability to correctly conceptualize his activities when he approaches the sling problem. This is so because the cognitive distress this conflict generates triggers the repression of the new conceptualization–that is, of the conceptualization the child frames to himself of his activities when throwing the ball with the use of a sling.

Piaget quickly recognizes that this hypothesis raises as many questions as it answers. For one thing, it is unclear what the nature of the 'conflict' that allegedly triggers the repression mechanism is. The purported conflict cannot be located–as Kitchener (1986, 64), for instance, claims–between (1) sensori-motor schemes at the unconscious level and (2) concepts at conscious level, because the application of unconscious schemes does not rely on (accurate) conceptualization. Hence, a child may employ a particular sensori-motor scheme at the unconscious level while simultaneously possessing an inaccurate conceptualization of this sensori-motor scheme without experiencing any inner conflict whatsoever–and certainly without experiencing the kind of inner conflict capable of triggering the repression (defense) mechanism.

The conflict also cannot be located at the purely conscious level–that is, between two different conceptualizations (concepts): one which is accurate and one which is inaccurate. In this case, according to Piaget, both concepts must by definition be conscious, since Piaget postulates that conceptualization and becoming conscious represent identical processes. This, however, is not the case, since the child is entirely unaware of the correct conceptualization of his activity when using the sling, and hence, does not possess contradictory conceptualizations (at the conscious level) (Piaget, 1974c/1976, p.340).

Piaget thus concludes: "It [the conflict] must therefore be situated in the process of conceptualization that characterizes the process of cognizance" (1974c/1976, p.341). Having identified the locus of conflict, Piaget may begin to sketch out a solution, which he introduces as follows :"[...]what precedes conception ('subception') and is defined as an 'unconscious perception' could well be accompanied by a certain consciousness at the moment it actually occurs, but this remains temporary in the sense that it does not seem to be integrated into the subsequent states" (Piaget 1974c/1976, 341f). According to Piaget, there exists a certain sub-phase of the process of construction (subception), during which the temporary conceptualization of a new concept occurs. If this new concept contradicts an existing concept, it will be fleshed out during the conceptualization process in only the barest of ways, and then immediately repressed. Piaget is thus compelled to introduce the notion of gradations of consciousness in order to disentangle the problems surrounding the repression mechanism.

But why, one may ask, are these (bare) new concepts immediately and entirely repressed? Why are the children unable to recall and to verbally report the experience of having entertained concepts that contradict prior conscious knowledge? As far as one can tell from Piaget's descriptions of the sling experiment, the children are aware of any conflict taking place. This is, indeed, strange, since it is a fact that there is a conflict between concepts (i.e. they logically contradict each other) and it is by definition a conflict between two *conscious* concepts.[9] Piaget's failure to adequately address this issue is likely to leave few readers satisfied with his

[9] Recall that, according to Piaget, conceptualization is itself the act of becoming conscious, so concepts must be conscious (at least before being repressed)

proposed solution, compelling them to seek answers elsewhere.

The crux of Piaget's confusion appears to lie in his identification of the process of conceptualization and the process of becoming conscious, since this identification forces Piaget to introduce spurious claims regarding a sophisticated repression mechanism in order to explain the much simpler fact that children may possess unconscious concepts. As we have seen, Whitehead does not follow this path. According to Whitehead, the formation of propositions (concepts) is not (necessarily) accompanied by consciousness. Consciousness arises in intellectual prehensions if and only if concept and reality do not correspond. Since at the purely propositional level there is no consciousness, it is quite possible for an individual to possess unconscious concepts (propositions) that contradict each other. Hence, for Whitehead, the conflict is, indeed, a conceptual conflict. However, it is located at the unconscious level. t The newly created concept is 'repressed' (i.e. ejected from consciousness) in favor of an already existing concept that has been confirmed on many previous occasions. Since a subject may hold concepts either consciously or unconsciously, conflicts between concepts may occur at the conscious or unconscious level. This allows Whitehead to provide a more plausible account of the repression mechanism and to sidestep the trap to which Piaget falls victim.

Piaget's argument is questionable at the very least. But is there any evidence in his body of research that may justify the introduction of the concept of gradations of consciousness? Piaget offers such a possibility in *Success and Understanding*: "[T]he negative qualities have no significance except in relation to other objects for comparison, to predictions, which can be used for a comparison, to predictions, that do not come true, or to momentary needs that are not satisfied. […] We are dealing here with a particularly important example for the *general progress of cognizance*" (Piaget, 1974a/1978, p.230f emphasis mine). This suggests that for Piaget, as for Whitehead, cases in which concepts fail to adequately represent the part of reality to which they refer stimulate the development of consciousness more actively.

Thesis 6: It is the function of consciousness to improve the re-establishment of balance

When human beings encounter changes in their environment, old cognitive patterns may be unable to assist them in coping with unfamiliar phenomena and situations. The mismatch between new external stimuli and old (conditioned) internal cognitive processes triggers cognitive disequilibrium in the individual. The individual's cognitive machinery seeks re-equilibration (i.e. the cognitive balancing of new information with old information) with the goal of establishing cognitive adjustments that complement new stimuli. Piaget states: "As for that equilibration process itself, in the course of this book we have frequently stressed the fact that transcendences of contradiction, in their double aspect of *extensional widening of the referentials* and a *qualitative transformation of notions in the direction of relativization*, are both active and constructive in nature" (1974b/1978, p.306 emphasis mine). Reconceptualization is thus a form of equilibration that generates improved concepts,–i. e. concepts that better enable the individual to respond to (new) environmental challenges. Through equilibration, the internal structure of the concepts (and their interrelations) becomes more differentiated and, as a consequence, the individual may adapt more successfully to his surroundings.

According to Piaget such 'equilibration majorant' is closely linked to the production of stability–that is, the more elaborate a cognitive system is, the more stable it is in the face of disturbing environmental influences (see for instance: 1976, p.371). A highly equilibrated cognitive system thus increases an organism's chance of survival. On this point, Whitehead and Piaget again differ. While Piaget believes the principle function of consciousness lies in improving an organism's fitness for survival, Whitehead is skeptical of this view. Consciousness is the source of potentially dangerous new ideas that may undermine an organism's (or an entire species') chance of survival (Whitehead 1933/1967, p.146). Hence, according to Whitehead, the principle function of consciousness is to intensify emotional experience and expression.

3. Final Comments

The above comparison of Whitehead's and Piaget's positions concerning a number of weak problems of consciousness has shown that both scholars share a wide range of convictions. The empirical evidence Piaget collects in his scientific research as the basis of his theoretical position thus confirms many aspects of Whitehead's theoretical position, as well.

Let us recall, however, that Whitehead and Piaget disagree on two fundamental issues: (1) whether conceptual conflicts can occur at an unconscious

level, and (2) what the principle function of consciousness is. In addressing the first issue, Whitehead's account is clearly the more convincing of the two. On the second point, the case is not so clear-cut. It appears that Whitehead's answer–that the principle function of consciousness is to elevate aesthetic and affective experience–gains the advantage only if one is convinced by Whitehead's criticism of (Neo)Darwinian evolution theory.

Throughout the course of this study, I have argued in favor of Whitehead's position. In an attempt to avoid repeating these points, I will end by instead discussing the implications of Whitehead's views as they relate to two closely connected topics: learning and consciousness (3.1) and artificial consciousness (3.2).

3.1 Consciousness and Learning

In most modern societies learning is mainly consciousness-based: in their lectures, teachers explicitly present the current state of the research of a given subject or sub-domain, stated in well-formulated sentences and packed into educational canons of knowledge that are often squeezed into text books, according to prescribed curricula for each school year. Students are then expected to memorize these contents and reproduce them on tests. One of the leading researchers in the domain of Artificial Intelligence (AI), Ron Sun, who has developed a hybrid cognitive architecture (CLARION), critically comments on this situation:

Most educational settings focus on teaching conceptual (explicit [conscious]) knowledge. While this may be appropriate for some subject areas, other subject areas may require learning information (e.g., features of complex systems or categories) that are better learned (at least initially) through extensive hands-on experience than with lectures or textbooks alone (that is with explicit learning alone). […] For example, to be a scientist is not simply a matter of accumulating a large amount of specific and explicit knowledge; it is, more importantly, a matter of developing intuitive (implicit) understanding and intuitive (implicit, procedural) skills for approaching and exploring scientific issues and for conducting scientific experiments (Sahdra & Thagard, 2003). […] In this light, understanding the interaction of implicit and explicit learning becomes a critical issue (Sun, Mathews & Lane, 2007, p.1).

Whitehead, too, is critical of the highly abstract form of traditional teaching and learning that pervades contemporary education in schools. According to Whitehead, learning under normal circumstances (consider, for example, the young child learning to speak (his native) natural language) starts at an unconscious level through direct interaction with, and experience of, the world.[10]. Whitehead calls the first of three learning phases that constitute a full learning cycle, a 'romance,' and criticizes traditional (i.e. abstract and linear) educational methods for failing to value and accommodate this phase while focusing solely on the phase he calls 'precision.' As Whitehead states, "It is evident that a stage of precision is barren without a previous stage of romance: *unless there are facts which have already been vaguely apprehended in their broad generality, the previous analysis is an analysis of nothing. It is simply a series of meaningless statements about bare facts, produced artificially and without any further relevance"* (Whitehead, 1929/1967, p.18 emphasis mine). Without prior, partially-unconscious interaction with a problem situation, students will find a teacher's academic instruction arbitrary, abstract, and ultimately meaningless.

Thomas Shultz, a leading scholar in the field of connectionist learning (a discipline inspired by simulations of human cognitive architectures in virtual machines)draws attention to the fact that connectionist learning research offers many recommendations for education, and particularly, for designing learning opportunities, that are backed by empirical data. One recommendation concerns the use of examples in teaching: "Teaching with examples is compatible with the classically important idea of learning by doing" (Shultz, 2007, p.1500). According to this approach, students should be provided with specific concrete tasks or problems to solve for themselves.

Schultz' second recommendation provides a complement to the first. According to Shultz, this second "important recommendation from connectionism is to accompany examples with *complete feedback about appropriate responses to each problem*. Full target feedback is more informative than the evaluative feedback provided by rewards which was emphasized in classical connectionism. Computational results show that converting information about being wrong into a useful target vector makes learning both slower and more difficult (Hertz et al., 1991)" (Shultz, 2007, p.1500 emphasis mine. Schultz clarifies that the data provided by connec-

[10] Whitehead's vision of this direct interaction is based on two primitive forms of perception that he postulates: (1) causal efficacy and (2) presentational immediacy

tionist learning research should not lead us to advocate a radical constructivist approach to learning, which eliminates teacher feedback and argues that students themselves construct new knowledge structures, but rather, should encourage us to adopt a moderate constructivist approach, which acknowledges that teachers (and their instruction) are important elements in the learning process.

In his final discussion of the lessons to be drawn from connectionist research, Shultz draws attention to the fact that "these recommendations coming from connectionist research (practice, feedback, prior knowledge, well-structured lessons) at first may appear more consistent with teacher-centered, rather than child-centered approaches to education. This seems a bit paradoxical given that constructivist approaches (such as CC and KBCC) are quite consistent with a Piagetian approach to knowledge acquisition that serves as psychological basis of much child-centered education. (2007, p.1501).

Shultz strongly emphasizes that "constructivist connectionist modeling suggests a rapprochement, by providing computational demonstrations that effective learning incorporates both of these approaches" (2007, p.1501). Whitehead, too, believes that effective learning is both student-centered and teacher-centered. Indeed, he posits a three-phased learning cycle that explicitly incorporates both approaches: the phases of romance (phase 1) and generalization (phase 3) are strongly child-centered, while phase 2, the so-called phase of precision, is teacher-centered.

In the first phase (romance), the child is presented with a well-selected problem which s/he should try to solve (but without any pressure to succeed); in the second phase (precision), the teacher introduces available knowledge relevant for solving the problem and connects this information as closely as possible to the students' learning experiences in phase 1, thereby ensuring that this information is familiar enough to be meaningful for the students. In the last phase (generalization), a return to romance occurs: the students again are presented with a new problem to solve.[11]

Whitehead's approach to education may be theoretically interesting, but, one may wonder, is it effective in practice? Having recently completed a two-year study investigating this and other related questions, I can confirm that Whitehead's approach leads to successful learning. During the course of this study, seven physics teachers provided academic instruction for students from grades five to eight (mean age: 12.4 years). Students in the treatment group received instruction consistent with Whitehead's three-phased learning cycle, while children in the control group received traditional (non-phased) academic instruction. Students who received Whiteheadian-style instruction for the full two years exhibited positive effects in the following areas: emotional wellbeing, academic self-concept, behavior during class, and overall cognitive development (Hagenauer, Strahl, Kriegseisen & Riffert, 2018; Riffert, Hagenauer, Kriegseisen & Strahl, 2020). Whitehead's ideas on unconscious and conscious processes have thus led to an instructional approach (learning cycles) which has been empirically tested and shown to provide positive results.

3.2 Artificial Consciousness – possible within Whitehead's metaphysics?

There does not appear to exist any axiom in Whitehead's metaphysical theory that, in principle, would preclude the possibility of machine consciousness. Since consciousness develops in actual entities, and since everything that exists–including artificially-created (manmade) machines–is constituted by actual entities, there is no reason to assume from the outset that machines could not support the development of consciousness. That being said, we should bear in mind that only complex actual entities can perform judgments (intellectual prehensions), and only judgments (particularly negative judgments) can have conscious subjective forms. Thus, for a machine to develop consciousness, it would have to be constructed in a way that allows it to harbor complex actual entities. Though this task appears daunting, it may, in fact, be achievable.

Hybrid computer architectures (HCA), which try to model natural human cognitive architectures, appear capable of producing intellectual prehensions (judgments); they consist of (1) a primitive, presymbolic form of (distributed) processing of external stimuli[12] and (2) a symbolic level of processing that builds on the results of the primitive processing level. HCA may produce results that are at variance with one another when they process data at both of these levels, providing an opportunity for the HCA to form 'intellectual judgments' (or something very similar) regarding the accuracy of these data. It,

[11] For a much more detailed presentation and discussion of Whitehead's approach to learning in light of contemporary empirical research, see: Riffert, 2005b

[12] Compare this to the unconscious causal efficacy of human beings that Whitehead describes

therefore, seems reasonable to suppose that a machine could develop *conscious* prehensions–at least occasionally.

Machine consciousness, however, would prove very short lived if the machine developing conscious prehensions could not also ensure the development of a 'living personal society.' Living personal societies enable actual entities to store and retrieve conscious contents, which is a precondition of possessing conscious contents for any length of time. It is the living personal society that, for example, allows an individual to consciously reflect on a scientific problem for hours, days, or even years.

The environments in which HCA and other machines engage in information processing appear, prima facie, simpler than the complex and ever-evolving (given the fact of neuroplasticity) human brain. Therefore, while it is possible that machines might develop consciousness, it seems unlikely, given the current state of technology, that machines will develop 'long-lived' conscious states like those that characterize human thinking. Thus, from a Whiteheadian perspective, research in AI that aims at creating virtual machines capable of exhibiting human-like consciousness must focus on the development of serially ordered living personal societies that are able to stabilize conscious states. As Whitehead put it: "A complex inorganic system of interaction built up for the protection of the 'entirely living nexus', and the originative actions of the living elements are protective of the whole system. On the other hand, the reactions of the whole system provide the intimate environment required by the 'entirely living nexus'. We do not know of any living [and therefore human-like conscious] society devoid of its subservient apparatus of inorganic societies" (Whitehead ,1929/1978, p.103).

Within a Whiteheadian metaphysical framework, it is not the development of (machine) consciousness that is improbable, but rather the stabilization (and hence, persistence) of consciousness. This raises the following questions: (1) how could a machine's conscious knowledge be stabilized from moment to moment, allowing the machine to reflect on it for a longer period of time?, and (2) how could a machine's conscious knowledge be stored so as to allow access/retrieval after long periods during which the machine is not consciously focused on this knowledge?[13] Only by finding solutions to these questions could machine consciousness begin to approach anything resembling human consciousness.

Thus, ninety years after the release of Whitehead's magnum opus, *Process and Reality*, this book still "holds within itself unexplored connexions with possibilities half-disclosed by glimpses and half-concealed by the wealth of material" that may continue to guide and stimulate contemporary research (Whitehead, 1929/1967, p.17).

[13] Consider the fact that human beings can retrieve conscious knowledge after long periods of not being conscious at all–e.g., following sleep or comatose states

References

Aleksander I (2005) *The World in My Mind, My Mind in the World*. Exeter, England: Imprint Academy.

Aleksander I (2007) Modeling Consciousness in Virtual Computational Machines. Functionalism and Phenomenology. *Synthesis Philosophica* 44(2): 447–454.

Anderson J R (1983) *The Architecture of Cognition*. Cambridge, MA: Harvard University Press.

Baars B J (2007) The Global Workspace Theory of Consciousness. In *The Blackwell Companion to Consciousness*, ed. M. Velmans and S. Schneider, 236–246. Malden, MA: Blackwell.

Barrett L F, Tugade M M, Engle RW (2004) Individual Differences in Working Memory Capacity and Dual-process Theories of Mind. *Psychological Bulletin* 130: 553–573.

Berry D (1993) Neurological Evidence. In *Implicit Learning: Theoretical and Empirical Issues*, ed. D. Berry and Z. Dienes, 113–128. Mahawah, NJ: Erlbaum.

Broadbent D E (1954) The Role of Auditory Localization in Attention and Memory Span. *Journal of Experimental Psychology* 47: 191–196.

Bröderbauer S, Huemer M, Riffert F (2013) On the Effectiveness of Incidental Hints in Problem Solving. *Gestalt Theory* 35(4): 349–364.

Bunge M (1980) *The Mind-Body Problem. A Psycho-Biological Approach*. Elmsford, NY: Pergamon.

Cataldo P (1982) Whitehead and Aristotle on Propositions. *Process Studies* 12(1): 15–23.

Chaiken S, Trope Y (1999) *Dual Process Theories in Social Psychology*. New York, NY: Guilford Press.

Chalmers P (1994) Review of Journal of Consciousness Studies. *The Times Literary Supplement* Nov. 25: 22.

Chalmers D (1995a) Facing up to the Problem of Consciousness. *Journal of Consciousness Studies* 2: 200–219.

Chalmers D (1995b). The Puzzle of Conscious Experience. *Scientific American* 273: 80–86.

Chella A, Manzotti R. (Eds) (2007) *Artificial Consciousness*. Exeter, England: Imprint Academy.

Cherry E C (1953) Some Experiments on the Recognition of Speech, with One and with Two Ears. *Journal of the Acoustical Society of America* 25(5): 975–979.

Claparéde E (1905/1911) *Experimental Pedagogy and the Psychology of the Child*. New York, NY: Longmans.

Damasio A (1994) *Descartes' Error*. New York, NY: Putnam.

Devine P (1989) Stereotypes and Prejudice: their Automatic and Controlled Components. *Journal of Personality and Social Psychology* 58: 5–18.

Dienes Z, Perner J (1999) A Theory of Implicit and Explicit Knowledge. *Behavioral and Brain Sciences* 22: 735–808.

Dijksterhuis A, Aarts H (2010) Goals, Attention and (Un)Consciousness. *Annual Review of Psychology* 61: 467–490.

Epstein S, Pacini R (1999). Some Basic Issues Regarding Dual Process Theories from the Perspective of Cognitive-experiential Theory. In *Dual-Process Theories in Social Psychology* ed. S Chaiken and Y Trope, 462–482. New York, NY: Guilford Press.

Evans J St B T (2009) How Many Dual-process Theories Do We Need? One, Two or Many? In *In Two Minds: Dual Processes and Beyond*, ed. J St B T Evans and K Frankish, 31–54. Oxford, England: Oxford University Press.

Evans J St B T (2012) Dual Process Theories of Reasoning: Facts and Fallacies. In *The Oxford Handbook of Thinking and Reasoning*, ed. K Holyoak and G Morrison, 115–133. Oxford, England: Oxford University Press.

Fetz R L (1984) Zur Genese ontologischer Begriffe: Für eine Verbindung Whiteheadscher und Piagetscher Ansätze. In *Whitehead und der Prozessbegriff/Whitehead and the Idea of Process*, ed. H Holz and E Wolf-Gazo, 409–434. Freiburg, Germany: Alber.

Fetz R L (1999) Whitehead, Cassirer, Piaget. Drei Denker – ein gemeinsames Paradigma. *Salzburger Theologische Zeitschrift* 3: 154–168.

Frith C, Rees G (2007) A Brief History of the Scientific Approach to the Study of Consciousness. In *The Blackwell Companion to Consciousness*, ed. M. Velmans and S. Schneider, 9–21. Malden, MA: Blackwell.

Fu Q, Bin G, Dienes Z, Fu X, Gao X (2013) Learning Without Consciously Knowing: Evi-dence from Event-related Potentials in Sequence Learning. *Consciousness and Cognition*, 22: 22–34.

Gazzaniga M S (1970) *The Bisected Brain*. New York, NY: Appleton-Century-Crofts.

Griffin D (1998) *Unsnarling the World-Knot. Consciousness, Freedom, and the Mind-Body Problem*. Berkeley, CA: The University of California Press.

Hagenauer G, Strahl A, Kriegseisen J, Riffert F (2018) Emotionen von Schülerinnen und Schülern im Physikunterricht auf Basis des Lernzyklenunterrichts. Befunde einer zweijährigen Interventionsstudie [Students' Emotions in Physics Classes during Learning Cycle Instruction: Results from a Two-year Intervention Study]. In *Emotionen und Emotionsregulation in Schule und Hochschule* [*Emotions and Emotion Regulation in School and University*], ed. G Hagenauer and T Hascher, 303–322. Göttingen, Germany: Waxman.

Kahneman D (2011) *Thinking: Fast and Slow*. New York, NY: Farrar, Straus & Giroux.

Kelley G A (1955) *The Psychology of Personal Constructs* (2 Vols.). New York, NY: Norton.

Kokinov B, Yoveva ? (1996) Context Effects on Problem Solving. In Proceedings of the 18th Annual Conference of the Cognitive Science Society, ed. G. W. Cottrell, 586–590. Hillsdale, NJ: Erlbaum.

Kokinov B (1997) Micro-level Hybridization in the Cognitive Architecture DUAL. In *Connectionist-symbolic Integration: From Unified to Hybrid Architectures*, ed. R Sun and F Alexander, 197–208. Hillsdale, NJ: Lawrence Erlbaum.

Kokinov B, Petrov A (2001) Integration of Memory and Reasoning in Analogy-making: The AMBR-Model. In *The Analogical Mind: Perspectives from Cognitive Science*, ed. D Gentner, K Holyoak and B Kokinov, 59–124. Cambridge, MA: MIT Press.

Kragh U (1955) *The Actual-Genetic Model of Person-Perception*. Lund, Sweden: Leerup.

Krashe S (1981) *Second Language Acquisition and Second Language Learning*. London, England: Pergamon.

LeDoux J (1992) Brain Mechanisms of Emotion and Emotional Learning. *Current Opinion in Neurobiology* 2(2): 191–197.

Lieberman M D (2003) Reflective and Reflexive Judgment Processes: A Social Cognitive Neuroscience Approach. In *Social Judgments: Implicit and Explicit Processes*, ed. J P Forgas, K R Williams and W von Hippel, 44–67. New York, NY: Cambridge University Press.

Lieberman M D (2007) The X- and the C-systems: The Neuronal Basis of Automatic and Controlled Social Cognition. In *Social Neuroscience*, ed. E Harmon-Jones and P Winkleman, 290–315. New York, NY: Guilford Press.

Marcel A (1983) Conscious and Unconscious Perception: Experiments on Visual Masking and Word Recognition. *Cognitive Psychology* 15(2): 197–237.

Milner A D, Goodale M A (1995) *The Visual Brain in Action*. Oxford, England: Oxford University Press.

Milner A D, Goodale M A (2008) Two Visual Systems Re-viewed. *Neuropsychologia*, 46: 774–785.

MaysW (1977) *Whitehead's Philosophy of Science and Metaphysics*. The Hague, Netherlands: Martinus Nijhoff.

Nagel T (1974) What is it like to be a Bat? *The Philosophical Review* 83(4): 435–450.

Nagel T (1987) *What Does it All Mean?* Oxford, England: Oxford University Press.

Norman J (2002) Two visual systems and two theories of perception: An attempt to reconcile the constructivist and ecological approaches. Behavioral and Brain Sciences, 25: 73–96; discussion: 96–144.

Oeser E, Seitlberger F (1988). *Gehirn, Bewusstsein und Erkenntnis*. Darmstadt, Germany: Wissenschaftliche Buchgemeinschaft.

Piaget, J. (1918). *Recherché*. Lausanne, Switzerland: La Concorde.

Piaget J (1924/1969) Judgment and Reasoning in the Child. Ottawa, Canada: Littlefield Adams.

Piaget J (1936) *La naissance de l'intelligence chez l'enfant*. Neuchatel, Switzerland: Delachaux & Niestle.

Piaget J (1937) *La construction du réel chez l'enfant*. Neuchatel, Switzerland: Delachaux et Niestle.

Piaget J (1950). *Introduction à l'épistémologie génétique*. Paris, France: Presses Universitaires de France.

Piaget J (1954) The Problem of Consciousness in Child Psychology: Developmental Changes in Awareness. In *Problems of consciousness*, ed. H A Abramson, 136–177. New York, NY: Josiah Macy Jr. Foundation.

Piaget J (1970) Inconscient affectif et inconscient cognitif. *Raison Presente* 19: 11–20.

Piaget J (1972) Erkenntnistheorie der Wissenschaften vom Menschen. Frankfurt a. M., Germany: Ullstein.

Piaget J (1973) *Der Strukturalismus*. Freiburg, Switzerland: Walter.

Piaget J (1974a). *Reussir et comprendre*. Paris, France: Presses Universitaires de France. (English: (1978) *Success and Understanding*. London, England: Routledge and Kegan.)

Piaget J (1974b) *Recherches sur la contradiction*. Paris, France: Presses Universitaires de France. (English: (1980) *Experiments in Contradiction*. Chicago, Il: University of Chicago Press.)

Piaget J (1974c) *La prise de la conscience*. Paris, France: Presses Universitaires de France. (English: (1976) *The Grasp of Consciousness*. Cambridge, MA: Harvard University Press.)

Piaget J (1974d) *The child and reality*. London: Muller.

Piaget J (1975). L'néquilibration des structures cognitives: problème central du développement. Paris, France: Presses Universitaires de France.

Piaget J (1976) *Die Äquilibration der kognitiven Strukturen*. Stuttgart, Germany: Klett.

Piaget J (1983) *Biologie und Erkenntnis*. Frankfurt a. M., Germany: Fischer.

Pons F, Harris P L (2001). Piaget's Conception of the Development of Consciousness: An Examination of Two Hypotheses. *Human Development* 44: 220–227.

Pons F, Harris P L, Rosnay M (2012) Piaget and Consciousness: Retrospect and Prospect. In *After Piaget*, ed. Eduardo Marti and Cintia Rodrigues, 99–122. New York, NY: Transaction Publishers.

Popper K R (1966) *Of Clocks and Clouds*. St. Louis, MO: Washington University Press.

Popper K R (1974) *Objektive Erkenntnis – Ein evolutionärer Entwurf*. Hamburg, Germany: Hoffmann & Campe.

Popper K R, Eccles J (1977) *The Self and Its Brain: An Argument for Interactionism*. Berlin, Germany: Springer.

Pothast U (1987) Etwas über Bewusstsein. In *Theorie der Subjektivität*, ed. Konrad Cramer, H Fulda, Rolf-Peter Horstmann and Ulrich Pothast, 15–43. Frankfurt a. M., Germany: Suhrkamp.

Reber A S (1967) Implicit Learning of Artificial Grammar. *Journal of Verbal Learning and Verbal Behavior* 6: 855–863.

Reber A S (1993) *Implicit Learning and Tacit Knowledge: An Essay on the Cognitive Unconscious*. Oxford, England: Oxford University Press.

Reverberi C, Toraldo A, D'Agostini S, Skrap M (2005) Better Without (Lateral) Frontal Cortex? Insight Problems Solved by Frontal Patients. *Brain* 128: 2882–2890.

Riffert F (1995) *Whitehead und Piaget. Zur Interdisziplinären Relevanz der Prozessphilosophie* [*Whitehead and Piaget. On the Interdisciplinary Relevance of Process Philosophy*]. Vienna, Austria: Lang.

Riffert F (2002) On Non-Substantialism in Psychology – Convergences between Whitehead's Process Philosophy and Piaget's Genetic Structuralism. *International Journal of Field-Being* 2(1) (Online Journal, no pagination) available at: http://www.fieldbeing.org//doc/Riffert_Non-Substantialism%20in%20Psychology.pdf [access: April 29, 2019)

Riffert F (2003) On Scientific Confirmation of Causal Efficacy. In *Searching for New Contrasts. Whiteheadian Contributions to Contemporary Challenges in Neurophysiology, Psychology, Psychotherapy, and the Philosophy of Mind*, ed. F Riffert and M Weber, 153–180. New York, NY: Lang.

Riffert F (Ed.) (2005a). *Alfred North Whitehead on Learning and Education: Theory and Application*. Newcastle, England: Cambridge Scholar Press.

Riffert F (2005b) Whitehead's Cyclic Theory of Learning and Contemporary Empirical Educational Research. In *Alfred North Whitehead on Learning and Education: Theory and Application*, ed. F Riffert, 89–119. Newcastle, England: Cambridge Scholar Press.

Riffert F (2012) Analytic Philosophy, Whitehead, and Theory Construction. *Process Studies* 41(2): 235–260.

Riffert F, Bröderbauer S, Huemer M (2015) Some Reflections on the Relation between Whitehead's Process Philosophy and Gestalt Psychology. *Process Studies* 44(2): 164–189.

Riffert F, Hagenauer G, Kriegseisen J, Strahl A. (2020) On the Impact of Learning Cycle Teaching on Austrian High School Students' Emotions, Academic Self-Concept, Cognition, and Behavior in Science Education. *Research in Science Education* (paper submitted).

Roy D (2007) *The Correlation of the Modes of Perception and the Phases Concrescence in Whitehead's Cosmology*. In *Perception Reconsidered. The Process Point of View*, ed. F. Riffert, 19–54. New York, NY: Lang.

Ryle G (1943) *The Concept of Mind*. London, England: Hutchinson.

Schacter D L (1987) Implicit Memory: History and Current Status. *Journal of Experimental Psychology Learning Memory and Cognition* 13(3): 501–518.

Schacter D L, Chiu C-Y P, Ochsner K N (1993).Implicit Memory: A Selective Review. *Annual Review of Neuroscience* 16(1):159–182.

Schmidt R (1994). Implicit Learning and Cognitive Unconscious: Of Artificial Grammars and SLA. In *Implicit and Explicit Learning of Languages*, ed. N Ellis, 165–209. London, England: Academia Press.

Seager W (2003) Whitehead and the Revival (?) of Panpsychism. In *Searching for New Contrasts. Whiteheadian Contributions to Contemporary Challenges in Neurophysiology, Psychology, Psychotherapy, and the Philosophy of Mind*, ed. F. Riffert and M. Weber, 333–345. New York, NY: Lang.

Seth A (2009) The Strength of Weak Artificial Consciousness. *International Journal of Machine Consciousness* 1(1): 71–82.

Sherburne D (1966) *A Key to Whitehead's Process and Reality*. Chicago, Il: University of Chicago Press.

Shultz T R (2007) Connectionist Learning and Education: Applications and Challenges. In Proceedings of the Twenty-ninth Annual Conference of the Cognitive Science Society, ed. D McNamara and G Trafton, 1497–1502. Mahwah, NJ: Lawrence Erlbaum.

Sloman S A (1996) The Empirical Case for Two Systems of Reasoning. Psychological Bulletin 119: 3–22.

Smith E R, Collins E C (2009) Dual-process Models: A Social Psychological Perspective. In *In Two Minds: Dual Processes and Beyond*, ed. J St B T Evans and K Frankish, 197–216. Oxford, England: Oxford University Press.

Smith E R, DeCoster J (2000) Dual Process Models in Social and Cognitive Psychology: Conceptual Integration and Links to Underlying Memory Systems. *Personality and Social Psychology Review* 4: 108–131.

Smith G (1957) Visual Perception – An Event Over Time. *Psychological Review* 64: 306–313.

Smith G (2004) *The Process Approach to Personality. Perceptgenesis and Kindred Approaches in Focus*. New York, NY: Springer.

Smolensky P, Legendre G (2006) *The Harmonic Mind. From Neuronal Computation to Optimality-Theoretic Grammar*. 2 Vols. Cambridge, MA: MIT Press.

Stoltz T (2018) Consciousness in Piaget: Possibilities of Understanding. *Psicologia: Reflexao e Critica* 31: 2–9.

Strawson G (2006) Realistic Monism: Why Physicalism Entails Panpsychism. *Journal of Consciousness Studies* 13(10–11): 3–31.

Sun R (1999) Accounting for the Computational Basis of Consciousness: A Connectionist Approach. *Consciousness and Cognition* 8: 529–565.

Sun R (2001) Computation, Reduction, and Teleology of Consciousness. *Journal of Cognitive Systems Research* 1: 241–249.

Sun R (2016) *Anatomy of the Mind: Exploring Psychological Mechanisms and Processes with the Clarion Cognitive Architecture*. New York, NY: Oxford University Press.

Sun R, Slusarz P, Terry C (2005) The Interaction of the Explicit and Implicit in Skill Learning: A Dual Process Approach. *Psychological Review* 112: 159–192.

Sun R, Mathews R C, Lane S M (2007) Implicit and Explicit Processes in the Development of Cognitive Skills: A Theoretical Interpretation with Some Practical Implications for Science Education. In *Educational Psychology Research Focus*, ed. E Vargios, 1–26. Hauppauge, NY: Nova Science Publishers.

Toates F (2006) A Model of Hierarchy of Behavior, Cognition and Consciousness. *Consciousness and Cognition* 15: 75–118.

Velmans M, Schneider S (Eds) (2007) *The Blackwell Companion to Consciousness*. Malden, MA: Blackwell.

Weiskrantz L (1986) *Blind Sight. A Case Study and Implications*. Oxford, England: Oxford University Press.

Weiskrantz L (1997) *Consciousness Lost and Found*. Oxford, England: Oxford University Press.

Whitehead A N (1925/1967) *The Function of Reason*. New York, NY: Free Press.

Whitehead A N (1929/1978) *Process and Reality*. (Corrected Edition) New York, NY: Free Press.

Whitehead A N (1933/1967) *Adventures of Ideas*. New York, NY: Free Press.

BALKAN JOURNAL OF PHILOSOPHY. Vol. 11, Issue 2, 2019

CONCEPTUAL ANALYSIS IN THE PHILOSOPHY OF SCIENCE*

Martin Zach

Department of Philosophy and Religious Studies, Faculty of Arts,
Charles University, nám. Jana Palacha 2, 116 38 Prague, Czech Republic
Department of Analytic Philosophy, Institute of Philosophy, Czech Academy of Sciences,
Jilská 352/1, 110 00 Prague, Czech Republic
email: m_zach@seznam.cz / zach@flu.cas.cz

Abstract:
Conceptual analysis as a method of inquiry has long enjoyed popularity in analytic philosophy, including the philosophy of science. In this article I offer a perspective on the ways in which the method of conceptual analysis has been used, and distinguish two broad kinds, namely philosophical and empirical conceptual analysis. In so doing I outline a historical trend in which non-naturalized approaches to conceptual analysis are being replaced by a variety of naturalized approaches. I outline the basic characteristics of these approaches with illustrative examples, arguing that recent developments in the philosophy of science show that in order to achieve a more adequate understanding of scientific endeavour we need to prioritize the naturalized accounts of the method.

Key words: conceptual analysis; empirical philosophy of science; naturalism.

1. Introduction

The philosophy of science established itself as an academic discipline in the first half of the twentieth century. At its inception the way in which the philosophy of science was practiced was highly influenced by the developments of that time. In particular, the logical analysis of the language of science was thought to be the key to understanding the structure of scientific theories, the bearing of evidence on a proposed hypothesis, the role of experiment, and the explanations provided by scientific theories. As in other fields of analytic philosophy, philosophers of science have also relied on conceptual analysis while analysing issues such as the above. In this article I offer a perspective on the ways in which the method of conceptual analysis has been used, and distinguish two broad kinds: philosophical and empirical conceptual analysis. In so doing I outline a historical trend in which non-naturalized approaches to conceptual analysis are being replaced by a variety of naturalized approaches.

I will characterize philosophical conceptual analysis by way of drawing attention to three key features: (i) analysing meta-scientific concepts independently of science; (ii) the unjustified tendency to use methods of formalization, and (iii) the requirement to analyse concepts in terms of necessary and sufficient conditions. Philosophical conceptual analysis is best seen in the schools of logical positivism and empiricism, which ruled for decades until they began losing influence under the weight of emerging internal issues prone to increasingly severe criticism. Despite the introduction of numerous attempted modifications intended to resolve the problems (see, e.g., Suppe, 1974 for an exhaustive overview), the project was ultimately abandoned. However, we may find traces of the logico-positivistic approach to conceptual analysis in schools that were formed as an antidote to the positivistic picture of science, as well as in some contemporary schools and debates. That said, the trend seems to be such that newly emergent approaches to the study of science are gaining in popularity. I categorize these other approaches, jointly referred to as 'empirical conceptual analysis', into four kinds: the case-study approach, applied approaches, qualitative approaches, and quantitative approaches.

* I am indebted to Lilia Gurova for providing me with much valuable feedback on a draft of this paper. I also wish to thank Vlasta Sikimic for discussing selected parts of this paper, and the audience at SILFS 2017 where an earlier version was presented.

As stated, the purpose of this paper is to characterize the method of conceptual analysis used in the philosophy of science. Ultimately, I will argue that we should prioritize the empirical ways of using this method if the goal is to achieve a better grasp of scientific endeavour. I will substantiate these claims by showing the strengths and weaknesses of these methods– both in comparison with one another and with traditional approaches. I will also illustrate some of the gains that have been made using empirical conceptual analysis, which might otherwise have been impossible to achieve through the application of traditional methods..

The structure of this article is as follows. Section 2 introduces philosophical and empirical conceptual analysis as two general accounts (i.e. non-naturalized and naturalized), followed by the provision of more detailed characteristics and illustrative examples (Sections 3 and 5). Section 4 serves as a link between the two extremes, illustrating the existence of combined approaches, thus supporting my view that it is not necessarily a question of either/or, but rather that middle-ground positions may and in fact do exist. Despite this fact, in Section 6, I present evidence that suggests we would do best to cut down on the non-naturalized aspects of the method of conceptual analysis, and to commit more fully to naturalized versions instead.

A final introductory remark concerns the way in which I shall retain a tight, specific focus. There is an ever-growing literature on the value of conceptual analysis in the context of the emergence of 'experimental philosophy' (see Knobe and Shaun, 2017 for an overview). I have little to add to it and will restrain myself from further comment. I believe this is justified, since my aim here is different from that which those engaged in experimental philosophy have in mind. Usually, what is at stake there is the use of an intuition-driven justification of philosophical thought experiments.[1] Furthermore, it is often couched in terms of the *a priori* and *a posteriori* distinction, with the experimental philosopher objecting to *a priori* methods of justification. Although there are some implicit similarities with some of the issues discussed here, my concerns are somewhat different and do not necessarily address the justificatory use of intuitions, nor the context of thought experiments.

Because of space constraints I also cannot discuss the Canberra plan[2] in any detail. The only remark I wish to make is that the Canberra plan is, in my view, appropriately placed in the non-naturalized camp because of its top-down nature and because it exhibits all three features of philosophical conceptual analysis (see below).

2. Two approaches to conceptual analysis in the philosophy of science

There is little disagreement on the point that philosophers employ conceptual analysis as a method of inquiry; however, it remains controversial what this method actually comes down to. Here I attempt to distinguish two broad accounts of the method, which I call *philosophical conceptual analysis* and *empirical conceptual analysis* respectively. I associate the former with a non-naturalized approach, whereas the latter is introduced as naturalized. In Section 3, I go into greater depth on this topic to further illuminate and distinguish these categories. Before going into more detail, let me first make a couple of general remarks on what I take conceptual analysis to be. Conceptual analysis is understood as the analysis of concepts in terms of other concepts, and any reflection on concepts can be construed as conceptual analysis.[3] However, the exact methods by which this analysis of concepts proceeds are manifold.

2.1. Philosophical conceptual analysis as a non-naturalized approach

Traditionally–exceptions notwithstanding–philosophical inquiry has aimed at analysing concepts in the form of addressing the question, 'What is X?'. In the context of the philosophy of science, by 'X' one usually means a meta-scientific concept such as *science*, *scientific explanation*, *law of nature*, *theory*, *model*, *experiment*, *values in science*, *scientific reasoning*, *reductionism*, etc. The goal of such an inquiry is to specify conditions under which a given concept applies to a given case. This means that in order for something to count as a genuine case of, for instance, scientific explanation, it has to satisfy the specified criteria: otherwise it is not a

[1] This is in no way exhaustive of the concerns of experimental philosophy, see, e.g., Sytsma and Buckwalter (2016).

[2] See Papineau (2015, Section 2.3) for a brief overview and a source of references.

[3] Thus, not only philosophy but also science heavily relies on such analysis. Generally, any introduction to a scientific discipline is focused on conceptual understanding and is often high-yield at the start.

case of scientific explanation. For example, according to the deductive-nomological account of scientific explanation, an explanation must meet several criteria: the explanandum must be a logical consequence (in the strict logical sense of the term) of the explanans, the explanans must contain general laws which must be required for the derivation of the explanandum, and the sentences constituting the explanans and explanandum must be true (Hempel and Oppenheim, 1948; see also, e.g., Godfrey-Smith, 2003, ch. 13 for discussion). This implicitly presupposes a certain methodological approach to studying concepts–one that, as I will argue, proves to be problematic.

Such an approach takes a concept as a starting point, usually with considerable philosophical baggage, and proceeds to define this concept in terms of various (potentially specious) normative criteria. A long-favoured assumption (with little justification) was to think of scientific explanations in terms of deriving conclusions (statements about observations or empirical laws) from more fundamental laws.[4] Thus, in analysing concepts, the starting point is often located within an inherited theoretical background against which analysis takes place, and as a result, a meta-scientific notion is investigated through a philosophical prism. Such an approach can be construed as a top-down approach: the philosophical prism sits at the top while a philosopher peeks down through it at the science.[5] One danger lurking in the shadows as a consequence of this approach is that science may end up distorted in ways that fit preconceived philosophical categories. I call this approach *philosophical conceptual analysis* to reflect the metaphor of the philosophical prism.

2.2. Empirical conceptual analysis as a naturalized approach

The other broad approach may also target the question, 'What is X?' The difference is thus not necessarily one of subject matter, but of the way in which the question is addressed. Rather than first proposing criteria for the applicability of a concept (with an intent of locating anything that seems to satisfy these criteria within contemporary scientific practice), *empirical conceptual analysis* takes as its starting point the concept as it already is employed within science. Hence, the nature of the approach is bottom-up: a philosopher studies what scientists take to be paradigmatic examples of explanations and proceeds to offer careful generalizations based on these, thus forming (philosophical) accounts of scientific explanation. The notion of the 'empirical' in 'empirical conceptual analysis' serves to suggest that the starting point is empirical, rather than philosophical in nature. The main point here is that conceptual analysis cannot be detached from studying the actual uses of concepts (and the ways in which we learn concepts). One may worry that, in this way, philosophical inquiry is reduced to a merely descriptive endeavour, whereby philosophy must sacrifice any claim to the normative dimension of a theory of science. It is beyond the scope of this paper to sufficiently address this objection; however, a few remarks are in order.

First of all, the idea of reducing the philosophy of science to a merely descriptive project is not necessarily a self-undermining enterprise, since the philosophy of science would still have its place among disciplines that take science as their subject matter. This is because, among other things, scientific disciplines are in the business of doing science rather than providing elaborate 'meta-scientific' theories. Furthermore, disciplines such as the sociology of science or the psychology of science focus on different aspects of scientific practice, i.e. the social dimension of science, including the institutional foundations of the research, funding, and publication system, and the cognitive dimension of engaging in scientific inquiry, respectively.[6] Second, it is not the case that the philosophy of science would thereby give up on any normative claims. Rather, the relationship would be much more complex. I will further discuss empirical conceptual analysis in Section 5, describing in more detail various empirical methods and examples.

2.3. Two caveats

Before I proceed any further, a couple of caveats are in order. The two broad approaches to con-

[4] The idea that the laws of nature play an indisputable role in explanation was questioned early on (Scriven, 1962), though for the most part in the context of ordinary explanations. It took some time before philosophers became aware of various explanatory practices within science which do not rely on the laws of nature in any substantial sense.

[5] A common worry raised by numerous authors is that such an approach is an exercise in the rational reconstruction of science, which leads to spurious results.

[6] This is not to suggest that there is no possibility of fruitful collaboration between the disciplines. Indeed, it very well may be that the contrary is true. The point is only that, on average, philosophers tend to focus on different aspects from sociologists or psychologists.

ceptual analysis specified above are not incompatible, strictly speaking. One may rely heavily upon one or the other, but one may also employ a combination of the two (see Section 4). This means that the two approaches ought to be construed as approaches that allow for various degrees of use. For example, a particular analysis may be predominantly based on the first approach while also somewhat drawing on the second.

From what I have said thus far one may be under the impression that both the top-down and bottom-up approaches are straightforward. This is not my claim, and it is important to further elaborate this point. Just as there are degrees to which one may rely on either of these approaches, there are varying degrees of strictness with which an inquiry follows either the top-down or bottom-up approach. For instance, it would be rather naïve to construe the bottom-up approach as a simple matter of extrapolation, i.e. broadly generalizing from a small set of cases. Yet, piecemeal, step-by-step generalizations stemming from the *detailed* analysis of several cases seem promising, as long as one is careful enough to draw constrained and limited inferences. There has been awareness of this issue for some time (e.g., Shapere, 1987), but to what extent it is feasible, and which modifications in the process of generalizing are required, remain matters of controversy, indeed the focus of an ongoing debate (see, e.g., Chang, 2012; Wagenknecht, Nersessian, and Andersen, 2015b). Thus, I implore the reader to approach the notions of top-down and bottom-up as useful shortcuts for the purposes of this discussion–rather than as (1) matters of simple deduction from a general theory to a particular case, or (2) induction from a small set of examples to a general theory, respectively.

3. Characterizing philosophical conceptual analysis

Three main features of philosophical conceptual analysis shall be detailed and illustrated through examples: the analysis of meta-scientific concepts independently of scientific practice; heavy investment in formalizing procedures without any clear justification for so doing; and the requirement to provide necessary and sufficient conditions for defining concepts.

3.1. The analysis of meta-scientific concepts independently of science

Meta-scientific concepts usually designate high-level concepts which scientists implicitly employ. Examples of such concepts are easy to find, and include the notion of science and its demarcation from non-science and pseudo-science, the notion of scientific explanation, the notion of the laws of nature, and so on. Scientists intuitively use these notions, usually without explicitly providing a precise characterization or explanation. One of the goals of philosophical inquiry is to render explicit what is implicit: this includes elaborating on the nature of the aforementioned concepts. The question now is how to proceed.

An especially popular approach in certain circles is to think of these concepts as something that can be investigated relatively independently of actual scientific practices. This stems from the worry that were we to take into account the actual practices of individual scientists we would end up with a subjective view of science. We are told that a proper philosophical analysis provides an epistemological grounding for meta-scientific concepts, and such an analysis needs to be free from subjective and historically contingent aspects. After all, epistemology, as opposed to psychology, is about objective relations, or so the argument goes (Schickore, 2018).[7] The practice of proceeding in this way has often relied on describing everyday examples, accompanied by only a handful of introductory science textbook examples. The limited array of examples and their inability to represent larger, more complex explanatory strategies should give pause to the extent to which one may apply the results. That said, the outputs of such analysis have often been presented in an overly generalized manner, suggesting that this is how science in general works. For instance, this is how the work of Hempel and Oppenheim (1948) on scientific explanation has been criticized by some (Scriven, 1962).

The view that an ‘understanding of what science is (or does) must be independent of the content of any specific scientific beliefs’ (Shapere, 1987, p. 7) turns out to be problematic, and has been criticized

[7] This is an echo of the famous distinction between the context of justification and the context of discovery (Reichenbach, 1938). According to this distinction, philosophical analysis, when properly conducted, belongs to the justificatory realm, whereas actual practices in which particular subjects engage are deemed psychological, and thus not a target for philosophical inquiry.

by Shapere himself, among others. The main issue concerns the likelihood of distorting the picture of what science is, what it does, and how it does do whatever it does. It is not clear whether we can sensibly characterize what is, for instance, a scientific explanation if we do not investigate various advanced instances of scientists giving explanations. Rather, what is more likely is that instead of developing an account of scientific explanation, we start with what we take to be an explanation, based on a source other than scientific practice (e.g., ordinary everyday examples, superficial introductory textbook examples, etc.). However, there is no guarantee that this preconceived account of explanation has something interesting to tell us about scientific explanations. It might turn out that we have fallen prey to using a philosophical construct that purports to offer 'the account of scientific explanation'. However, such a designation is arbitrary because, in fact, it is an account of scientific explanation only in name, not substance.

As suggested, the independence of the analysis may not be absolute. Indeed, as illustrated above by Hempel and Oppenheim, analysis often progresses through a process of browsing scientific examples. One may thus wonder whether this is an instance of a bottom-up, rather than top-down approach. There are a couple of things to note. First, recall from the previous section that these approaches are to be construed as a matter of degree; hence, it may be possible to view such attempts as a bottom-up approach to some small degree (because the depth of these examples is usually rather shallow). Second, notwithstanding the first point, it still fits better with the top-down picture (to some degree) given the skewed picture such an approach often provides. Browsing through selected examples in a rapid fashion runs the risk of falling prey to the (philosophical) prism by which one has been previously influenced.[8] Thus, the advantages of the top-down approach may, in the end, outweigh those offered by the bottom-up approach.

3.2. Tendencies in formalizations

Valued by many, formalization has been construed as a tool that is put to work to help with clarifying various concepts. Formalization has received considerable attention in the context of the development of modern logic. Turning sentences of natural language into logical formulae has successfully eliminated many of the ambiguities in ordinary language, resulting in end-products that exhibit clear structure, and thereby provide greater insight. Similarly, many of the scientific disciplines that have proven extremely successful are steeped in mathematical formalisms of various kinds. The success of formalization has had a significant impact within science and logic (philosophy) , leading some to think that, perhaps, formalization may be the backbone of critical thought. Although formalization is a tool which has the undeniable capacity to clarify certain areas of inquiry, it is not self-sufficient, and as always, one has to carefully evaluate the merits and limits of using this method.

Whenever various means of formalization have been introduced into the social sciences, many have voiced their dissent.[9] Perhaps the most pressing issue concerns the question of whether by mimicking some of the features of well-established scientific disciplines we maintain reasonable epistemic standards for good practice–e.g. we avoid importing form that is devoid of meaningful content. Indeed, with the expansion of formalization we may run the risk of engaging in cargo cult science rather than actual science. A striking example illustrating some such issues is the 'second law of criminal behaviour', discussed in some length by Haack (2003). Originally proposed by David Abrahamsen, the law states that a criminal act (C) is equal to the sum of personal criminal tendencies (T) plus one's total situation (S), divided by the amount of one's resistance (R). Thus, we end up with a neat equation in the form of C = (T+S) / R. Such an equation seemingly provides a way to quantify an aspect of human behaviour. However, it is far from clear how to make sense of the equation, since the quantified notions lack both the necessary rigour (e.g. the notions of 'personal criminal tendencies' and 'one's total situation,' etc., remain vague) and a method for measuring them. Arguably, the social sciences have spent considerable time discussing such methodological pitfalls.

[8] This risk is perhaps ever present, included in any strong bottom-up leanings; again, the difference lies in the degree to which one is exposed to such a risk.

[9] Numerous reasons have been given in support of the claim that formalization (mathematization) may introduce unforeseen problems. It is outside the scope of this paper to enter into these debates, or even to provide an overview. However, in no way do I want to leave the impression that formalizing the social sciences is a fool's errand. Quite the contrary, I am committed to the view that various approaches to formalization have proven extremely fruitful within the social sciences, and do not believe this claim is undermined by particular attempts that have been misguided.

Coming back to the philosophy of science, even a brief overview of the literature provides examples of similar practices dangerously close to cargo cult science. A case in point is Larry Laudan's notion of the 'problem-solving effectiveness of a theory', which, according to him, 'is determined by assessing the number and importance of the empirical problems which the theory solves and deducting therefrom the number and importance of the anomalies and conceptual problems which the theory generates' (Laudan, 1978, p. 68). Since no precise way of counting the 'empirical problems', 'anomalies', and 'conceptual problems', nor any way of establishing a measure of the 'importance' of these problems have been provided, Laudan may be charged with engaging in contentious attempts to provide rigour where none is to be found. Although this is different from employing straightforward formalism, it nonetheless shares the rudiments.

Yet another related problem is found in a recent paper by Fiora Salis. In a debate over metaphysics and the semantics of fictions in the context of scientific models she argues that,

> when we say that Zaphod is more narcissistic than Morris we say (or imply) that there are certain degrees of narcissism i and j such that i > j, Zaphod has narcissism to degree i, and Morris has narcissism to degree j

and she continues by claiming that

> this is problematic because Zaphod and Morris do not exist [Zaphod and Morris are fictional characters], so they cannot really have degrees of narcissism i and j. In other words, Zaphod and Morris do not instantiate the relevant degrees of properties (Salis 2016, p. 248).

This problem, Salis suggests, is solved by an appeal to mathematical entities in a way that the above statement now takes the following form: "There are some degrees of narcissism, i, j, such that i > j, according to *The Hitchhiker's Guide to the Galaxy* Zaphod has narcissism to degree i, and according to *Changing Places* Morris has narcissism to degree j" (Salis, 2016, p. 249). Here the problem with formalization somewhat differs from the two cases discussed above, since in this case we already understand what the relevant concepts are (i.e. narcissism), and, perhaps, we also have an intuitive understanding of how to 'measure' these (this last point is contentious, but let us grant it here for the sake of argument). Arguably, such formalization does not advance our knowledge because it provides no new information beyond the ideas it recapilates. To be fair, Salis' debate is about metaphysics, and so I should say more about how it relates to the deeper issue I have chosen to examine. In Salis' argument, and in similar discussions, I contend that formalization plays a misleading role and distracts us from discussing an important issue related to the human cognitive power of imagination. Rather than encouraging reflection, such formalizationt presents us with a problem in metaphysics that has spurious means of resolution.[10] Thus, formalizationserves merely as a deceptive intermediary, diverting our attention from a respectable problem to one which is questionable.

Finally, Hempel and Oppenheim (1948) went to great pains to provide a formalization of what they perceived as the logical structure of scientific explanation. Couched in the language of first order logic, they believed their account could provide much needed rigour. In doing so, they relied on the strict logical notion of derivation. However, as noted by Teller (2001), among others, the way in which derivation in science (even in highly mathematized fields, such as physics) often proceeds is a far cry from the notion of derivation familiar to logicians.. It is hard to see how an account of scientific explanation could be based on such a strict logical notion of derivation if actual explanatory practices are not.[11]

I hasten to add that my quarrel here is not with the methods of formalization. Again, it is an undeniable fact that various formal methods play a key role in many of the sciences. Nor do I wish to claim that certain phenomena cannot, by their nature, ever be approached from the perspective of some formal method, including those addressed in the philosophy of science. Rather, the claim here concerns only the way in which some of these formal methods have been put to work. I argue that in many instances the attempts of philosophers are questionable, insofar as

[10] Note that appeals to mathematical realism are considered by many to be highly contentious (Field, 1989).

[11] There are many more examples of spurious formalization tendencies. For instance, an attempt by Lewis (2001) to construe the pessimistic induction argument in the debate on scientific realism in terms of a base rate of true and false theories has been criticized by Saatsi (2005). More specifically, Lewis' view that the characterization of the reliability of a test in the context of this debate should be construed in terms of the rates of false positives and false negatives has been the target of criticism.

they provide little justification for choosing and using that particular method.[12] Indeed, it may very well be that using formalization to account for certain aspects of science is an inadequate way to conduct an inquiry, and promoting formalization in these instances would be ill-advised. In other words, we must first consider whether employing a formal method may help us to learn something useful, and then make sure the method is accompanied by content, so that we avoid falling into the trap of cargo cult research.

3.3. The search for necessary and sufficient conditions

The third feature of the philosophical conceptual analysis discussed here is the requirement to provide necessary and sufficient conditions for defining concepts. Thus, to understand a concept according to philosophical conceptual analysis is to possess a set of necessary conditions that are jointly sufficient for applying a given concept. Meta-scientific concepts should then be defined in terms of those conditions that specify whether the concept genuinely applies to a purported instance of that concept. For instance, according to the deductive-nomological (DN) account of scientific explanation, a scientific explanation consists of deriving whatever is to be explained from the premises, i.e. from a set of laws of nature and initial conditions. Furthermore, the premises must be true, so that the derivation is not only valid, but also sound. This set of premises plus the requirement of truth provides the defining feature of what it means to have a scientific explanation of a given phenomenon.

The driving force behind this idea is that having a set of necessary and sufficient conditions will provide a demarcation of concepts, so that we may achieve precision and clarity in applying them. This is echoed throughout the history of the philosophy of science and is favoured by many to this day. Consider the debate on scientific representation. Gabriele Contessa claims that "according to the interpretational conception, a vehicle is an epistemic representation of a certain target (for a certain user) if and only if the user adopts an interpretation of the vehicle in terms of the target" (Contessa ,2007, p. 57). Analogously, Adam Toon says that

> theories of depiction aim to state conditions that are necessary and sufficient for something to be a depiction. Similarly, if possible, we want to provide a set of conditions that are both individually necessary and jointly sufficient to establish an instance of each form of model-representation that we identify (Toon, 2012, p. 23).

The sentiment is shared by Roman Frigg and James Nguyen:

> the problem is: what turns equations and structures, or fictional scenarios and physical objects into representations of something beyond themselves? It has become customary to phrase this problem in terms of necessary and sufficient conditions and throughout this review we shall follow suit (Frigg and Nguyen, 2017, p. 51).

Because I return to this issue in Section 5.2, I shall not comment any further here.

3.4. Summary

Let us now summarize what has been said thus far. In thinking about various aspects of science, philosophers employ the method of conceptual analysis. I first provided a rough distinction between two ways of construing this method: philosophical conceptual analysis and empirical philosophical analysis. The former can be characterized by three key features: the analysis of meta-scientific concepts being pursued independently of actual science, it often relies on (1) unjustified formalization and it seeks to (2) define concepts in terms of a set of necessary and sufficient conditions, i.e. in an essentialist manner. For these reasons, philosophical conceptual analysis may be construed as a (3) non-naturalized methodology, i.e. one that is not properly informed by science. Before providing the characteristic features of empirical conceptual analysis, I shall consider approaches that stand somewhere in between, which draw on both the philosophical and empirical aspects.

4. Combining philosophical and empirical conceptual analysis: semi-naturalized approaches

Earlier I noted that philosophical and empirical conceptual analysis may both be practiced to various

[12] For instance, think of the 'raven paradox' discussed in the literature on confirmation theory. It proved misguided to analyse confirmation by using the apparatus of predicate logic. That does not mean, however, that all formal approaches are doomed to fail, as, for instance, some Bayesians have argued. Indeed, Bayesian statistics is a tool that is used in confirmatory practices in some domains, including medicine.

degrees. That would suggest the possibility of identifying approaches that lie somewhere between the two extremes. Since I use the terms 'naturalized' and 'non-naturalized' to refer to positions on the opposite ends of a spectrum, it is appropriate to call those that are somewhere in-between 'semi-naturalized'. I shall now briefly discuss two such semi-naturalized approaches: the historical school of the philosophy of science, as exemplified by the works of Thomas Kuhn, and the social constructivist movement of Bruno Latour and Steve Woolgar.

Thomas Kuhn's *The Structure of Scientific Revolutions*, originally published in 1962, is now considered a classic. Kuhn's success lay in the novelty of his approach and in his engaging prose. He had both enthusiastic supporters and sworn critics. In the second edition of *The Structure*, Kuhn (1970) attempted to address some of the most pressing issues surrounding his notion of incommensurability, which had proven troublesome and hard to swallow for the many who accused Kuhn of e relativizing science to the extent that science would no longer enjoy the status of a privileged way of knowing. According to these detractors, Kuhn had failed; he subsequently spent the bulk of his career trying to develop his ideas into a non-relativistic philosophy of science.

Although Kuhn was not the first to realize that psychological and sociological factors play an important role in everyday scientific practice, he arguably had the most profound and catalyzing impact on the ensuing development of this idea. Indeed, a significant portion of the philosophy and sociology of science is informed by Kuhn. While Kuhn's recognition of the psychosocial aspects of scientific practice is consistent with models that are taken for granted in the more empirically-minded circles of the philosophy of science, it was widely rejected during the heyday of logical positivism.

However, as noted above, Kuhn has additionally been criticized by a number of commentators (see, e.g., Shapere, 1964) for clinging to certain aspects of logical positivism. One of the most controversial notions Kuhn introduces is the thesis of incommensurability, which states roughly that scientists working within different paradigms (both historical and contemporary) cannot understand each other; they employ different vocabularies, and even when they use the same word, the word has a different meaning for each of them. Kuhn's incommensurability thesis relies on the descriptivist[13] theory of reference a philosophical account of how terms allegedly refer. However, he entirely fails to justify his choice of preferred philosophical theory. As noted by Shapere (1964), a bizarre consequence of the incommensurability thesis (based on the descriptivist theory of reference) is that physicists belonging to different 'paradigms' could not possibly understand each other, something that is clearly not the case. Not only can physicists understand each other with respect to current competing theories, they can also understand their colleagues and rivals across time, i.e. they understand, or make sense of, most of the physics that was developed long before their time.[14]

Thus, it seems that Kuhn fell prey to the philosophical prejudice that biased his construction of a theory of science. There is no a priori reason why the philosophical theory of reference should be privileged over observation when considering the question of whether scientists from different paradigms can understand each other.[15] , Kuhn is consequently guilty of engaging in philosophical conceptual analysis, insofar as he analyzes meta-scientific concepts independently of empirical scientific methods. I should quickly add that this does not define the whole of his work, but only certain aspects, as his notion of incommensurability arguably illustrates. Indeed, Kuhn's work represents a mixture of approaches: on the one hand, he clearly demonstrates his leaning toward a more empirical construc-

[13] According to this account, the reference of a term is fixed by its descriptions. If the descriptions change then the referent changes as well. To pick one of Kuhn's examples, consider the concept of mass. In Newtonian physics, mass is an intrinsic property of objects and as such stays invariant. However, in the Special Theory of Relativity, mass actually acquires another meaning, one that is dependent on the relative velocity of the object (STR effectively distinguishing between 'rest mass' and 'relativistic mass'). According to this view, when scientists from the two paradigms speak about mass they mean completely different things.

[14] Of course, no one means to suggest that, say, a theoretical physicist can understand every detail of what an experimental physicist does, or vice versa. Nor can a theoretical physicist in one particular field understand every detail of another theoretical physicist's work from a different field. However, the reason here is different from what Kuhn suggests, i.e. it is a lack of familiarity with domain-specific details that inhibits understanding, rather than an inability to communicate ideas due to the lack of a shared (scientific) discourse.

[15] Note that what is at dispute is understanding each other rather than agreeing with each other.

tion of a philosophical theory by providing historical case studies and by acknowledging the role of psychological and sociological factors in actual scientific practice; on the other hand, his theory is ensconced in a prior philosophical framework that distorts certain of its aspects.

The second example concerns the social constructivist school of the philosophy of science, the strong programme in the sociology of knowledge (Latour and Woolgar, 1979). Latour's and Woolgar's *Laboratory Life* shows great promise in introducing empirical methods into the philosophy of science – in their case, the ethnographic method.[16] I will not delve into the prospects of the ethnographic method here, but will note that it may offer helpful insight into the daily business of scientific research, among other things, by examining what scientists do and how they interact in lab meetings and in other contexts. Despite the potential merits of the ethnographic method, it is necessary to acknowledge the philosophical prejudice that skewed Latour and Woolgar's observations.

Latour and Woolgar observed the scientific process through the philosophical prism of relativism, a theoretical lens that was neither derived from, nor justified by their observations (which would represent an empirically-minded bottom-up approach). Latour's and Woolgar's work represents an instance of forcing science into a pre-established philosophical framework–in this particular case, the framework of relativism (a top-down approach). Intuitively, such a relativistic stance seems foreign to the basic principles of science. While an intuitive appeal is far from a decisive argument, it is enough to raise doubts about the strong conclusions of relativism. These doubts serve to indicate that Latour and Woolgar must provide a stronger, evidence-based argument in order to demonstrate convincingly that scientific practice is, indeed, steeped in relativism. One may object that the truth of relativism can be established on merits other than what the scientific community thinks, which is surely true. However, in this case, one would arrive at relativism wholly independently of science, and it is unclear what source of knowledge could demonstrate the truth of relativism, p especially given the radical nature of this claim. Thus, the position of relativism seems to be nothing more than a philosophical construct.

[16] See Section 5.3 for more on the results the ethnographic method has brought to contemporary research.

5. Characterizing empirical conceptual analysis

Historically speaking, there appears to be an ever-growing movement among scholars that is characterized by a clear tendency to profess an inclination to a naturalistic approach toward the philosophy of science–in other words, an approach that is grounded in empirical knowledge. The precise nature of this approach, and the kind of empirical grounding that is required, however, remain contentious issues that philosophers have slowly started to address in greater detail. Below, I offer a brief characterization of several different approaches that seem to be gaining in popularity, with a special focus on what these methods of inquiry uncover.

In Section 2.2, I discussed empirical conceptual analysis in terms of a bottom-up approach. The starting point of such analysis is a concept as it is used in the sciences. The hope is that, by analysing the context in which this and similar cases emerge, we may find ourselves in a position to say something more substantive about that particular notion, and perhaps clarify it by so doing. I contend that there are four main branches of empirical conceptual analysis, each based on a different method. These include the case-study approach (Section 5.1), the applied approach (Section 5.2), and the qualitative (Section 5.3) and quantitative approaches (Section 5.4).[17]

The reason for ordering the discussion in this way is purely historical, i.e. philosophers first paid attention to scientific practices as exhibited in experimental research or as discussed in scientific papers, and only later have they espoused methods of a qualitative and quantitative nature.[18]

In the following sections I will provide an overview of several examples to further illustrate all four kinds of methods mentioned above and to show the points of departure from the traditional analyses as discussed in Section 3.

5.1. The case-study approach

The case-study approach aims to analyse partic-

[17] In scientific disciplines the case study approach is usually conceived as an instance of a qualitative method (Berg, 2001, ch. 10). In the philosophy of science, however, I believe it makes sense to keep them separate for historical reasons. In any case, nothing of substance hangs on this point.

[18] This is not to say that such a development is straightforward. Indeed, as we have seen in Section 4, some used qualitative methods early on. However, such few exceptions do not pose a threat to what is otherwise a strong trend.

ular examples of scientific practice, or particular historical periods, from the perspective of the history and philosophy of science; as such, it differs from an approach of merely using examples to illustrate one's point. While examples are often superficial and do not take up more than a paragraph or two, case studies consist of much deeper analysis and offer much deeper insight. Based on this distinction, it should be clear that I have illustrated my claims throughout this paper by enumerating examples, rather than examining case studies. Arguably, in his most celebrated work, Kuhn has also relied on examples, rather than case studies. Now, it is important to note that the distinction I am drawing is not determined solely on the basis of length, nor on an individual's acquaintance with the subject matter being discussed. The crux of this distinction lies instead in the following insight:, the case-study approach, notably, is consistent with the bottom-up approach discussed above (rather than the top-down approach). Indeed, according to the case-study approach, it should (for the most part) be particular fact-based cases that drive philosophical theorizing, and not the other way around. These fact-based cases are not limited to the subject matter of published scientific papers; the empirical realm of investigation includes scientific practices and findings derived from various sources, including one's own observations and experience.

The case-study approach is embodied by the 'new experimentalism' movement that started in the 1980s, notably in the works of Ian Hacking, Allan Franklin, Peter Galison, Robert Ackermann, and Deborah Mayo, among others.[19] Philosophers of this strand revolted against what they perceived as a theory-centric view of science that had become popular during that time. According to the popular theory-centric view, scientific research is principally a matter of formulating general theories in terms of sets of propositions, and experiment functions only to test the validity of a theory. The new experimentalists parted ways with the received view once they started investigating actual experimental practices, with the intent of providing accounts of what was taking place in working laboratories. Soon enough, these philosophers re-conceptualized the notion of experiment in terms of real-world experimental practices and the roles experiment plays in scientific practice generally. They started reporting on the equipment used, maintained, and calibrated in various sorts of labs; on the data-gathering processes, data-analysis processes, and the practices of intervening in a system; on cross-checking results, controlling for error, and the elimination of noise. Thus, thanks to the 'new experimentalism,' experiment has been re-conceived as something that "has a life of its own" (Hacking, 1983) and is often independent of a high-level theory. Indeed, experiment often functions differently from how tradition views it. For instance, experiment has an important exploratory function, something akin to plain old curiosity that may open up a completely new field of research. By removing the noise, experimental manipulation is also key to creating new phenomena and, as such, is most visible within the sphere of modern technology. Without paying attention to actual experimental practices, it would have been difficult for the new experimentalists to show the extent to which the received view provided only a very limited account of the nature of experiment.

Beside the new experimentalism movement, the 'new mechanism' movement has gained in prominence since the early 2000s.[20] The roots of the new mechanism movement stretch as far back as the early 1990s, most notably to the publication of Bechtel and Richardson's *Discovering Complexity*, originally published in 1993 (Bechtel and Richardson, 2010). It arose in the context of the study of cellular and molecular biology and soon spread to neuroscience, cognitive science, and the social sciences. In introductory remarks such as, "the account of mechanism (…) was motivated by particular examples of research in the biological sciences" we are again reminded that the method of inquiry is of the bottom-up sort, and that the starting point is scientific practice itself (Bechtel and Richardson, 2010, p. xx).[21] Research in many fields can plausibly be construed as a search for the underlying mechanisms responsible for observed phenomena, as evidenced by the actual usage of the notion of mechanism by scientists (Craver and Tabery, 2015). This is not to

[19] See especially Hacking (1983).

[20] See Machamer, Darden, and Craver (2000) for an early paper. A point of note is that this paper is the most cited paper in the recent philosophy of science, with over 2,000 citations. This further hints at the fact that the tides are changing, and such approaches are gaining in strength. See also Glennan and Illari (2018) for an authoritative overview of the cutting edge of the movement.

[21] The notion of 'example' here is intended to be used interchangeably with 'case study'. That said, elsewhere Bechtel and Richardson do use the notion of 'example' in the same way as discussed in the present article, and they distinguish it from 'case study'.

say that the concept of mechanism has an agreed upon meaning, nor that any single definition is sufficient to capture it. Indeed, there is no need for an exhaustive definition, since different disciplines and research projects may function according to slightly different variations of the more general notion.[22] Instead, we are given a qualitative characterization of the notion as it applies in specific research contexts, illustrating the way in which the new mechanists have abandoned traditional approaches of studying concepts.

Clearly, both the new experimentalism and the new mechanism movements revolt against all three features of the philosophical conceptual analysis discussed in Section 3. Their intense focus on actual scientific practice suggests a method of theoretical analysis that tries to be faithful to the process and principles of science. Within each of these movements there is very little formalization to speak of, and the quest for defining concepts in terms of necessary and sufficient conditions is deemed unilluminating.

Construed this way, we may now see the difference between the case-study approach, as exhibited by the new experimentalism and new mechanism movements on the one hand, and the philosophical conceptual approach on the other hand. Recall the point from Section 3 when I briefly discussed Hempel and Oppenheim's approach. Their starting point consisted of mostly superficial examples, on the basis of which they drew far-reaching conclusions, which they subsequently formalized. I argued that, to a certain degree, such an approach is better viewed as top-down, rather than bottom-up. We may now appreciate the point in greater detail: the case-study approach concerns advanced practices. It, moreover, requires substantial analysis of these, which significantly affects the philosophical conclusions that are drawn.

5.2. The applied approach: the case of the cognitive-historical method

Having discussed the case-study approach, let us turn to what I call the 'applied approach'. By 'applied' I mean to suggest that philosophers draw heavily on existing results, especially from the cognitive sciences, in order to address a particular issue. Hence, it may be said that these philosophers apply empirical results from previous research to their own investigations. In practice, the applied approach is intertwined with the case-study approach.[23] Rather than presenting a general (and inevitably vague) analysis, I will briefly describe a particular instance of such an approach.

Nancy Nercessian has championed what she calls the 'cognitive-historical' method for investigating the ways in which certain scientific concepts have changed over time (Nercessian, 2008). As the name of the method suggests, it combines both cognitive and historical approaches. Let us look in some detail at what this method has to offer.

The traditional approach to studying concepts consisted of searching for the essential properties, that is, in writing down the set of necessary and sufficient conditions that define a particular concept (see Section 3.3). This tradition has been challenged on at least two fronts. First, Wittgenstein (1953) argued that it is difficult, if not downright impossible, to define a great number of concepts in terms of necessary and sufficient conditions. We are unsuccessful in defining concepts in this way because some concepts are inherently fuzzy, with boundaries that seem to withstand attempts at sharpening them. Instead, Wittgenstein coined the term 'family resemblance' in order to capture the network of characteristics that relate particular instances to a given concept. Second, some cognitive scientists have also taken issue with the tradition, basing their opposition on research data. Rosch and Mervis (1975) conclude that concepts tend to have a graded structure: there are typical cases of a given concept and there are also cases which are more or less similar to them. For instance, while robins seem to be a typical example of the concept of birds, penguins and ostriches, though still birds, are less obvious examples of this concept. In connection with this, one typical characteristic of birds is that they have the capacity to fly, but clearly not all of them can.[24] Given these findings, Nersessian argues for a more nuanced ap-

[22] This much has been well recognized, though there seems to be some value in using a generalized notion of a mechanism that captures most of the otherwise disparate usages (see Illari and Williamson, 2012; Glennan, 2017).

[23] In practice, the applied approach is aimed at a particular case. However, the reason for keeping the applied and the case study approaches distinct is that the case-study approach does not need to draw on evidence from a specific branch of the hard sciences–for example, the cognitive sciences– in order to investigate a given case.

[24] Results coming from the cognitive sciences have greatly improved theories of concept formation and theories of reasoning (e.g. bounded rationality). Kornblith (2002, 2014) is a good source for an overview of the literature on the research into concept formation, with a discussion of its philosophical significance.

proach to studying the historical development of concepts, such as the field concept in the works of Faraday, Maxwell, Lorentz, and Einstein. In the construction of the notion of a field, pictorial visualization and analogy played a key role in the works of Faraday and Maxwell, as Nercessian documents.. Nercessian's aim is thus to closely track all the small changes and the various cognitive tools that have aided in the construction of broad scientific concepts.[25]

In a similar fashion, Shapere notes that "what is important in trying to understand scientific concepts is, (…), the way they develop; and to this task traditional theories of meaning, and indeed the whole program of studying meanings, is a hindrance rather than a help" (Shapere, 1987, p. 34). He goes on to claim that

> philosophical theories of meaning, such as those which seek necessary and/or sufficient conditions of application of terms, cannot do justice to the development of such concepts as 'force', 'energy', 'field', 'electron', 'particle', or, indeed, any scientific concepts at all, including allegedly 'metascientific' ones. For in scientific cases, the emphasis is on developing concepts which are adequate, in the light of what we have learned, for thinking and talking about nature and for understanding it (Shapere, 1987, p. 10).

Nercessian thus refrains from succumbing to the necessary and sufficient camp. Scientific practice is of utmost importance to her, and she uses no, or very little, formalization.

5.3. Qualitative methods

I have argued that some philosophy of science movements turned away from the traditional, non-naturalized way of inquiry. The earliest attempts, eventually gaining prominence, were fuelled by case studies, and by developments in cognitive science. More recently, however, philosophers of science have begun drawing even more on various empirical ways of conducting inquiry.

A vast array of qualitative methods is at the disposal of philosophers of science. Before moving on, some methodological worries need to be addressed. Some have argued that qualitative methods are suspicious–best to be avoided–because they rely on subjective reports or other criteria that are far from the rigour of quantitative methods.[26] Although this objection does carry some weight, I believe it may be addressed by realizing the proper domain of application, the particular goals of inquiry, and the fact that every method, including quantitative, has both limits and merits. Ethnographic research may allow a deeper insight into the workings of a small community; however, the price one has to pay is the highly constrained possibility of generating any broader (and broadly meaningful) generalization. Quantitative methods, e.g., survey methods, on the other hand, may allow generalization, but surveys cannot provide the kind of intimate insight made possible by ethnographic research.

The most popular methods used in the philosophy of science include ethnographic research, (mostly semi-structured) interviews, textual analysis (including textbooks, published papers, lab notes, grant proposals, etc.), and some combination of all of these. To fully appreciate how these methods may contribute to our understanding of science, let us examine some examples.

Using the ethnographic method in combination with several others, Kevin Dunbar has analysed scientific causal reasoning during interactions among scientists in weekly lab meetings (see, e.g., Dunbar, 2002). The purpose of these meetings is to address and attempt to solve problems that arise during research, many of which concern unexpected findings. Dunbar claims that "scientists have a clear set of cognitive strategies for dealing with these findings" (Dunbar, 2002, p. 157) by first categorizing their findings with respect to their current knowledge, and second, drawing an analogy with a related experiment. This may eventually lead either to the realization that there was something wrong with the experiment, or that, perhaps, a new discovery has been made. Though, as Dunbar emphasizes, causal reasoning is multifaceted, the method of analogy enjoys some popularity. In this context, Dunbar claims that

[25] It should be noted that the 'prototype' approach to concepts favoured by Rosch and Mervis (1975) is not the only game in town, as there are a few contenders. This means that Nercessian's method may be also based on not as strong a footing as originally thought. However, for my purposes here nothing of substance depends upon whether the specific details of Nercessian's account are correct. Rather, my point is to illustrate that some such approach is necessary to account for the conceptual development of scientific and meta-scientific notions (see below), and that it is empirically-oriented in spirit.

[26] For more details see, for example, (Brysbaert and Rastle 2009, ch. 10).

> it is important to note that analogies based upon superficial features are very useful and help solve problems that the scientists frequently encounter. For example, if a scientist is obtaining uninterpretable data and another scientist makes an analogy to an experiment that worked in which the temperature was different, and suggests changing the temperature, this can potentially fix the problem. Many of the superficial analogies were of this type and solved the problem. These types of analogy are very frequent, but were rarely mentioned by the scientists in their interviews and never appeared in their lab books. When asked why they did not mention these analogies, the scientists thought that they were so obvious that they were not worth talking about. Despite the obviousness of these analogies, they are one of the most frequent cognitive workhorses of the scientific mind (Dunbar, 2002, p. 159).

Dunbar also reports on the comparative differences among Canadian, American, and Italian immunology labs. Although these labs were of similar size, used the same types of materials, cells, and equipment, and published in the same leading journals, and although members of these labs even attended the same conferences, the reasoning during lab meetings somewhat differed from lab to lab in accord with the norms of the wider culture (Dunbar 2002, p. 163).

Importantly, all these findings are impossible to uncover by any method other than the participatory, i.e. ethnographic method, because as Dunbar explicitly states, these things are omitted from scientists' published works and lab notes.

Thus, using the qualitative approach, Dunbar has enriched our understanding of the concept of scientific reasoning. He approached the question of 'What is X?', i.e., what is scientific reasoning, from the perspective of investigating how scientists do, in fact, reason. Dunbar does identify the traditional deductive and inductive inferential strategies, but what his analysis brings forth is the way in which such characterizations of scientific reasoning remain only partial, and must be complemented by a whole array of other reasoning practices. Accordingly,we must enrich our notion of scientific reasoning.

Similarly, Susann Wagenknecht has combined the empirical methods of participatory observation and interview to investigate the ways in which scientific knowledge is generated in interdisciplinary labs (Wagenknecht, 2016). Her aim was to uncover how practitioners handle the 'epistemic dependence' on one another, and how 'epistemic trust' is established in the working environment of interdisciplinary teams,–more particularly, in a small planetary research group and a large molecular biology lab. The epistemic dependence stems from the fact that in collaborative efforts scientists with different backgrounds must rely on each other, since it is impossible for them to verify every detail of each other's research; this leads to the cognitive division of labour. Thus, trust is very important if researchers are to form a group that is based on epistemic dependence. Hence, these methods allow Wagenknecht to elaborate on the nature of the notions of epistemic dependence and epistemic trust.

Most recently, also drawing on the participatory observation ethnographic method, Rebecca Hardesty investigated a group of neurobiologists whose research focuses on the genetic underpinnings of Down syndrome (Hardesty, 2018). This group has developed a new mouse model ('GCDS') that they regard as 'the gold standard' of all Down's syndrome genetic research. However, there are severe practical limitations regarding the use of this model. These include the costs ($15,000 for each mouse), the short life-expectancy of the strain that is ill-suited for this particular research (i.e. age-related cognitive degeneration as it occurs), and the low chance of successful breeding (about 11%). Consequently, the lab uses a different mouse (TS65Dn) which they regard as inferior to their own. The goal of Hardesty's research was to answer questions such as, 'how does the lab justify working with what they believe to be an inferior mouse model?' (Hardesty, 2018, p. 15). In Hardesty's own words

> lab X has developed a new 'genetically correct' Down syndrome mouse model (GCDS) that it rarely uses. Instead, it uses a mouse which it considers to be an inadequate genetic model (TS65Dn) of the condition; however, the Lab has performed a series of experiments in order to show that the inferior Ts65Dn mouse is genetically equivalent with respect to APP in the GCDS mouse, therefore, Lab X has a new justification for the old mouse. Yet, Lab X has not revealed this equivalency explicitly in the papers that came out of these experiments; instead, they only published the data on APP in the Ts65Dn line (Hardesty, 2018, p. 20).

Furthermore, Hardesty argues that

> by solving this practical problem of not being able to use the mouse line and resolving this tension between their values of good practice, the lab made the GCDS mouse a standard against which one could evaluate other DS mouse models. This is an epistemic success for Lab X and a philosophically relevant feature of modeling practice that was not apparent without participation/observation fieldwork (Hardesty, 2018, p. 21).

The above is intended to demonstrate that the traditional philosophical conceptual method has limits with respect to its ability to illuminate important aspects of scientific practice. Furthermore, these examples[27] should also illustrate the way in which qualitative methods provide unique insight into scientific minds. Such insight may not be achievable by other methods, including other empirical methods, such as quantitative methods. This is because, as Dunbar notes, scientists may not mention all of the practices they rely on, including the kinds of analogies they use in their interviews or lab notes. Thus, qualitative methods may catch one's special attention only in participatory research. Let us now turn to quantitative methods.

5.4. Quantitative methods: the case of the OPTIMIST project

Quantitative methods are used to first collect data and then further analyse them using various statistical techniques. One form of the quantitative method, favoured in this type of empirical research in philosophy, is the survey. Participants are asked to fill in a questionnaire by indicating with which of the pre-defined answers they identify the most. Depending on the nature and purpose of the survey, the participants may also be asked about their gender, academic position, or other general information about themselves. This allows the gathering of a great deal of data in a standardized way, but by its design it requires some simplifications to be made, which may obscure seemingly minor details that later could turn out to be of some importance. This trade-off further illustrates the need to use a combination of methods to form an overall picture.

The OPTIMIST Project constitutes a recent example of employing quantitative research methods in the philosophy of science .[28] The acronym OPTIMIST stands for 'Optimization Methods in Science and Technology' and is a collaborative effort of researchers from the University of Belgrade and the Ruhr-University Bochum.[29] One branch of the project concerns itself with better understanding the motivations and preferences of scientists, and with gaining insight into their working conditions with respect to their environment and communication–all with the hope of providing some guidance in helping to maximize their efficiency. In order to address these issues, members of the project have launched a survey for researchers in laboratories such as Fermilab and the Rutherford Appleton Laboratory. This sort of reflection on values in science provides much-needed clarification of the concept of values and the various roles they play in scientific practice.

However, other quantitative methods, such as various data-mining methods (rather than surveys), have been put to work. Previously published results of the OPTIMIST project addressed the question of the connection between the efficiency of a series of high-energy physics experiments run in Fermilab, measured by publication and citation rates, and the team size, the time it takes to complete an experiment, and the number of teams per experiment (Perović, Radovanović, Sikimić, and Berber, 2016). Using data envelopment analysis, distance-based analysis, and analysing other contributing factors to team and project performance, the authors conclude that "[small teams] very simply proportionally widely outperform the bigger ones in terms of citations and publications" (Perović, Radovanović, Sikimić, and Berber, 2016, p. 102). Building on these results Sikimic, Radovanovic, and Perovic (2018) further address the question of the extent to which time spent on a particular project correlates with its outputs. They show that there comes an 'epistemic saturation point', after which the number of most significant results usually drops, whereas the number of less significant publications increases. They suggest that this may have a bearing on the rational organi-

[27] Some further examples can be found in a recent volume edited by Susann Wagenknecht, Hanne Andersen, and Nancy Nercessian, which contains a discussion of various qualitative methods as well as their application to particular research questions (Wagenknecht, Nersessian, and Andersen, 2015a).

[28] More precisely, the project is an interdisciplinary effort, and as such, includes researchers from disciplines such as computer science, psychology, science policy and the philosophy of science. More broadly, it also belongs to the field of 'social epistemology' (see below).

[29] For details visit http://www.ruhr-uni-bochum.de/optimist-survey/.

zation of research with respect to the decision of whether to continue a project or not.

A possible objection lurking behind is whether this sort of project still constitutes philosophy, an objection sometimes raised in debates on experimental philosophy. I have attempted to address this worry in Section 2.2, albeit in a limited fashion. I believe it is worth noting that the project, broadly construed, belongs to the field of social epistemology, which is an emerging field in philosophy, one that addresses the social nature of knowledge. As such, it is a legitimate field within philosophy (see, e.g., Goldman, and Blanchard, 2015 for an overview).[30]

6. Concluding remarks

A general distinction has been proposed between two ways in which conceptual analysis has been employed in the philosophy of science. On the one hand, I described philosophical conceptual analysis, exceedingly popular (not only) in the logical positivism movement, by reference to three characteristic features. These include, firstly, the conducting of analysis independently of the details of scientific practice. The purported justification is that by incorporating various actual details, the analysis of science would be in danger of succumbing to subjectivism, something which must be avoided in the philosophy of science, which concerns itself with the context of justification. Secondly, the notion that formalization is an adequate instrument for gaining insight into and clarifying meta-scientific concepts. Thirdly, that the analysis and genuine clarification of meta-scientific concepts requires providing a set of necessary and jointly sufficient conditions for the employment of a given concept.

This approach is partially inherited by movements that struggled to show the inadequacy of logical positivism, as I discussed in connection with Kuhn's historical school of the philosophy of science. Furthermore, we saw that traces of those features are still prevalent in some contemporary debates, including the debate on scientific representation. I argued that these features can be characterised as consistent with a non-naturalized methodology, because they do not pay due respect to actual scientific practice. However, as noted above, none of these features is inherently flawed. Rather, I argued that it is the way in which they have been used that is inappropriate.

The second general view of conceptual analysis, I suggested, is empirical. Empirical conceptual analysis is a bottom-up approach, rather than top-down; i.e. it takes scientific practice as its starting point. I described multiple methods employed by various movements and showed how they improve upon previous methods. Indeed, some of the interesting results would be difficult (if not impossible) to obtain through more traditional methods. This is one of the main reasons why philosophical conceptual analysis, even if practiced in a relatively modest way, (i.e. avoiding extremes, including a belief in the independence and complete objectivity of science, an overreliance on formalism, or an unhinged commitment to the 'necessary and sufficient approach' to concepts), is being slowly replaced by more empirically-oriented methods–as, indeed, it ought to be.

With the advent of new research methods and their introduction to the philosophy of science, many authors have begun to pay greater attention to how their approach differs from that of others. It is not a rare sight to lay one's eyes on passages in which philosophers voice critical opinions of non-naturalized methods. For instance, with respect to the debate on scientific explanation, Collin Rice claims that "unfortunately, (…) philosophers have often inappropriately attempted to apply their primarily *a priori* approaches to explanation to these unique instances in biology rather than first considering how these models are actually used by scientists to provide explanations" (Rice, 2015, p. 590-591). Similarly, with respect to the bottom-up and top-down approaches, the remarks of Bechtel and Richardson reflect many of the ideas laid out in this paper. For instance, they claim that

> it is worth noting that the history of science can be approached in two ways, as a reservoir of examples to illustrate a prior conception of science or as a source of insight into features of science that an adequate philosophical account of science needs to incorporate. Our use of the history of science is in the latter tradition. (…) We then built the account around the cases rather than selecting cases to illustrate the account. As we have said, our focus was on what counted as 'reductionistic' research in these fields. We soon discovered that these fields could not be naturally fitted to existing philosophical models of reduction

[30] This is in line with how Sikimic (personal communication) defends the philosophical aspect of the research.

> without doing serious damage to the sciences and to their actual historical development (Bechtel and Richardson, 2010, p. xx).

The form of naturalism exemplified by the empirical conceptual analysis discussed here is perhaps different from some of its other incarnations. For example, in some sense, this may be a stronger commitment to naturalism than found in some of Quine's work, as suggested by Bechtel, when he claims that

> although Quine made claims about scientific practice (e.g., about the role of nonepistemic factors such as simplicity in the evaluation of scientific theories) and about the limitations on scientific knowledge (e.g., the underdetermination of theories by all possible evidence), he generally did not engage science as a naturalist. For example, he did not study the investigatory strategies or reasoning employed by particular scientists or the specific explanations they advanced (Bechtel, 2008, p. 8).

Although, for reasons described above, I favour empirical conceptual analysis in general, I do not think philosophers should commit to one particular method. Every method provides access to a specific aspect, or set of aspects of scientific investigation, but not to others. Every method allows one to answer some sets of questions, but not others, as shown in Section 5, where I discussed the characteristics of and differences among the various methods, including the case-study approach, the applied approach, and qualitative and quantitative methods. In order to gain a complete picture of scientific practice, different methods of investigation must be used and the results then integrated.

The continuous shift toward empirical conceptual analysis as a method, or rather, a set of methods, is further evidenced by the establishment of international societies such as the Society for Philosophy of Science in Practice (SPSP) and the rate at which these communities have grown. We are also witnessing the emergence of collaborative and interdisciplinary projects such as the OPTIMIST project, or the Evidence-Based Medicine Plus movement (EBM+). The focus of these projects often extends beyond 'mere' analysis and has more practical objectives in mind. These include policy-making regarding, for example, the organization of research or clinical practice in medicine. Consequently, philosophers are now starting to jointly publish articles with scientists in both philosophical and scientific journals. All this suggests that we are witnessing a truly naturalistic shift with respect to the method of inquiry in the philosophy of science.

Let me close by addressing a legitimate worry one may have at this point. The literature on naturalism and various empirical methods in philosophy has become enormous. I have not addressed many of the pressing issues people have been discussing, but I would maintain this is because my purposes here target different aspects of the debate at large. It would prove impossible to address all that has been said on this theme. l Yet, I believe I have managed to bring focus to a particular problem within the philosophy of science in a way that furthers discussion and contributes to the growth and vitality of the discipline. I must admit that much of the discussion in the present article is sketchy at best and deserves further elaboration that is beyond the scope of this paper. I hope that such elaboration is made possible through discussion, debate, and meaningful engagement with the ideas I have proffered throughout this article.

Acknowledgement

This output was created within the project »Rationality Crisis and Modern Thought«, subproject »Principles of Naturalizing Philosophy of Science« realized at the Charles University, Faculty of Arts with financial support of the Specific university research in 2018.

References

Bechtel W (2008) *Mental Mechanisms: Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge.

Bechtel W, Richardson A (2010) *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research*. Cambridge (Mass.): MIT Press.

Berg B L (2001) *Qualitative Research Methods for the Social Sciences*. Boston: Allyn and Bacon.

Brysbaert M, Rastle K (2009) *Historical and Conceptual Issues in Psychology*. London: Prentice Hall.

Chang H (2012) Beyond Case-Studies: History as Philosophy. In *Integrating History and Philosophy of Science: Problems and Prospects*, eds Seymour Mauskopf and Tad Schmaltz, 109–124. Dordrecht: Springer.

Contessa G (2007) Scientific Representation,

Interpretation, and Surrogative Reasoning*. *Philosophy of Science* 74.1: 48–68.

Craver C, Tabery J (2015) Mechanism in Science. In *The Stanford Encyclopedia of Philosophy (Spring 2017 Edition)*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/spr2017/entries/science-mechanisms/.

Dunbar K N (2002) Understanding the Role of Cognition in Science: The Science as Category Framework. In *The Cognitive Basis of Science*, editeds Peter Carruthers, Stephen Stich, and Michael Siegal, 154–170. Cambridge: Cambridge University Press.

Field H (1989) *Realism, Mathematics, and Modality*. Oxford: Blackwell.

Frigg R, Nguyen J (2017) Models and Representation. In *Springer Handbook of Model-Based Science*, eds Lorenzo Magnani and Tommaso Bertolotti, 49–102. Dordrecht: Springer.

Glennan S (2017) *The New Mechanical Philosophy*. Oxford: Oxford University Press.

Glennan S, Illari P (eds) (2018) *The Routledge Handbook of Mechanisms and Mechanical Philosophy*. London: Routledge.

Godfrey-Smith P (2003) *Theory and Reality : An Introduction to the Philosophy of Science*. Chicago: University of Chicago Press.

Goldman A, Blanchard T (2015) Social Epistemology. In *The Stanford Encyclopedia of Philosophy (Summer 2018 Edition)*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/sum2018/entries/epistemology-social/.

Haack S (2003) *Defending Science - within Reason: Between Scientism and Cynicism*. Amherst: Prometheus Books.

Hacking I (1983) *Representing and Intervening*. Cambridge: Cambridge University Press.

Hardesty R A (2018) Much Ado about Mice: Standard-Setting in Model Organism Research. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences* 68–69: 15–24.

Hempel C G, Oppenheim P (1948) Studies in the Logic of Explanation. *Philosophy of Science* 15.2: 135–175.

Illari P M, Williamson J (2012) What Is a Mechanism? Thinking about Mechanisms across the Sciences. *European Journal for Philosophy of Science* 2.1: 119–135.

Knobe J, Shaun N (2017) "Experimental Philosophy." In *The Stanford Encyclopedia of Philosophy (Winter 2017 Edition)*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/win2017/entries/experimental-philosophy/.

Kornblith H (2002) *Knowledge and Its Place in Nature*. Oxford: Clarendon Press.

Kornblith H (2014) *On Reflection*. Oxford: Oxford University Press.

Kuhn T (1970) *The Structure of Scientific Revolutions*. Chicago: Chicago University Press.

Latour B, Woolgar S (1979) *Laboratory Life: The Construction of Scientific Facts*. Princeton: Princeton University Press.

Laudan L (1978) *Progress and Its Problems: Towards a Theory of Scientific Growth*. Los Angeles: University of California Press.

Lewis P J (2001) Why The Pessimistic Induction Is A Fallacy. *Synthese* 129.3: 371–380.

Machamer P, Darden L, Craver C F (2000) Thinking about Mechanisms. *Philosophy of Science* 67.1: 1–25.

Nercessian N J (2008) *Creating Scientific Concepts*. Cambridge (Mass.): MIT Press.

Papineau D (2015) Naturalism. In *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition)*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/win2016/entries/naturalism/.

Perović S, Radovanović S, Sikimić V, Berber A (2016) Optimal Research Team Composition: Data Envelopment Analysis of Fermilab Experiments. *Scientometrics* 108.1: 83–111.

Reichenbach H (1938) *Experience and Prediction*. Chicago: University of Chicago Press.

Rice C (2015) Moving Beyond Causes: Optimality Models and Scientific Explanation. *Noûs* 49.3: 589–615.

Rosch E, Mervis C B (1975) Family Resemblances: Studies in the Internal Structure of Categories. *Cognitive Psychology* 7.4: 573–605.

Saatsi J T (2005) On the Pessimistic Induction and Two Fallacies. *Philosophy of Science* 72.5: 1088–1098.

Salis F (2016) The Nature of Model-World Comparisons. *The Monist* 99.3: 243–259.

Schickore J (2018) Scientific Discovery. In *The Stanford Encyclopedia of Philosophy (Summer 2018 Edition)*, ed. Edward N. Zalta. https://plato.stanford.edu/archives/sum2018/entries/scientific-discovery/.

Scriven M (1962) Explanations, Predictions, and Laws." In *Minnesota Studies in the Philosophy of Science III*, edited by Herbert Feigl and Grover Maxwell, 170–230. Minneapolis: University of Minnesota Press.

Shapere D (1964) The Structure of Scientific Revolutions. *The Philosophical Review* 73.3: 383–394.

Shapere D (1987) Method in the Philosophy of Science and Epistemology: How to Inquire about Inquiry and Knowledge. In *The Process of Science: Contemporary Philosophical Approaches to Understanding Science*, ed. Nancy J. Nercessian, 1–40. Dordrecht: Martinus Nijhoff Publishers.

Sikimić V, Radovanović S, Perović S (2018) When Should We Stop Investing in a Scientific Project? The Halting Problem in Experimental. In *Proceedings of the XXIV Conference "Empirical Studies in Psychology,"* eds Kaja Damnjanović, Ivana Stepanović Ilić, and Slobodan Marković, 105–107. Belgrade: University of Belgrade.

Suppe F 1974() *The Structure of Scientific Theories.* University of Illinois Press.

Sytsma J, Buckwalter W (eds) (2016) *A Companion to Experimental Philosophy*. Malden, MA: Wiley-Blackwell.

Teller P (2001) Twilight of the Perfect Model Model. *Erkenntnis* 55.3: 393–415.

Toon A (2012) *Models as Make-Believe*. London: Palgrave Macmillan.

Wagenknecht S (2016) *A Social Epistemology of Research Groups*. London: Palgrave Macmillan.

Wagenknecht S, Nersessian N J, Andersen H (2015a) *Empirical Philosophy of Science : Introducing Qualitative Methods into Philosophy of Science*. Dordrecht: Springer.

Wagenknecht S, Nersessian N J, Andersen H (2015b) Empirical Philosophy of Science: Introducing Qualitative Methods into Philosophy of Science. In *Empirical Philosophy of Science: Introducing Qualitative Methods into Philosophy of Science*, eds Susann Wagenknecht, Nancy J. Nercessian, and Hanne Andersen, 1–10. Dordrecht: Springer.

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

BALKAN JOURNAL OF PHILOSOPHY. Vol. 11, Issue 2, 2019

NEWTON'S BUCKET (THOUGHT) EXPERIMENT

Maja Malec

University of Ljubljana, Faculty of Arts
Ljubljana, Slovenia
e-mail: maja.malec@ff.uni-lj.si

Abstract:
The bucket experiment in Newton's *Principia* is quite simple. Nonetheless, physicists as well as philosophers and historians of science are still debating its purpose and success. I present two interpretations found in the literature. According to the first, Newton tries to prove absolute rotation and thus the existence of absolute space. According to the second, he tries to provide a definition of absolute rotation as it is used in his mechanics. Closely connected to this is his rejection of Descartes' explanation of rotation and of motion in general. I conclude with a short discussion on whether the bucket experiment can be classified as a thought experiment.

Key words: Isaac Newton, space, motion, rotation, thought experiment, natural philosophy.

1. Introduction

Newton presents the bucket experiment in the introductory section of his *Principia*,[1] entitled "Definitions," where he defines key terms, such as mass, momentum, force, inertia, centripetal force, etc. More precisely, it is part of the "Scholium" added to the eight definitions, in which he discusses the concepts of time, space, place, and motion.

The experiment is quite simple. It involves a bucket, some water, and a long rope, and it takes Newton two thirds of a page to describe it and its results. Nonetheless, physicists as well as philosophers and historians of science are still debating its purpose and success. In this respect, it shares the fate of the whole "Scholium," that always puzzled commentators. For starters, was it written by Newton the physicist or Newton the philosopher? It also serves as a good basis for the consideration of the role that the concepts of space and time play in a physical theory, how the conceptions of space and time have developed from Newton to Einstein, how the philosophical conceptions relate to the physical ones, and what the relevance of the philosophy of space and time is in general. One can also ponder the methodological issues. Is Newton justified in making claims about absolute time, space, motion, rotation? Where is the limit of scientific theory, and how far is a scientist allowed to go in drawing metaphysical conclusions from the empirical findings of scientific experiments? How, indeed, ought one to distinguish between a well-grounded theoretical hypothesis and an unwarranted metaphysical speculation? Moreover, what if it is precisely metaphysical speculation that sparks a scientific breakthrough?

My aim in this article is quite modest. I will begin with Newton's own description of the example in the context of his view on time, space, and motion in the *Principia*, and then present two interpretations found in the literature on this topic. The first views the bucket experiment as Newton's attempt to prove the existence of absolute rotation and–especially–(metaphysical) absolute space from an empirical

[1] *Philosophiae naturalis principia mathematica* (*Mathematical Principles of Natural Philosophy*) was first published in 1687. The second edition appeared in 1713 and the third in 1726. I use the translation by Cohen and Whitman from 1999.

experiment, and the second views it as Newton's attempt to define a theoretical concept of absolute rotation within his physical theory. The latter, more charitable, interpretation is further supported if the bucket example is also recognized as a critique of Descartes' theory of motion, which can be generalized to all relativistic/relational theories of his time. I conclude with a short discussion on whether the bucket experiment can be classified as a thought experiment.

2. The bucket experiment in the context of Newton's views on time, space, and motion in the *Principia*

In the *Principia* Newton formulates classical mechanics and gravitational theory, providing a unified explanation of all natural phenomena. His mechanics is based on the absolutist view of time, space, and motion. The concepts of time, space, place, and motion are generally familiar to us and do not need to be defined, unlike technical terms,such as mass, force, or inertia, whose meaning warrants clarification The difficulty is that we come to know these concepts solely in relation to the objects of sense perception, which can affect our comprehension of them. Therefore, in the "Scholium" to "Definitions," Newton explains in what way these "quantities" are used in the *Principia*. He terms them as "absolute, true and mathematical" as opposed to the "relative, apparent and common" ones.[2]

Absolute time "in and of itself and of its own nature, without reference to anything external, flows uniformly,"[3] while relative time is any external measure of duration by means of some observed motion. Similarly, absolute space "of its own nature without reference to anything external, always remains homogeneous and immovable,"[4] while relative space is some movable measure of absolute space that is determined by the senses with respect to some arbitrary system of bodies. Similarly, the place of the body is the space it occupies, either absolute or relative, depending on the nature of the occupied space. And finally, absolute motion is "the change of position of a body from one absolute place to another,"[5] whereas relative motion is from one relative place to another. The example of a sailor moving in one direction on the deck of a ship that is traveling in the opposite direction illustrates how we can differentiate many relative motions depending on the chosen reference point. However, there is only one absolute motion in relation to absolute, immovable space.

The example is quite instructive, but the fact of the matter is we cannot know whether our ultimate reference point is indeed the absolute space or some relative space that is in motion relative to the absolute space. This we cannot know because space itself is empirically inaccessible, causally inert, or, as Newton puts it, "the parts of that immovable space in which the bodies truly move make no impression on the senses."[6] The same is true of the uniformly flowing absolute time. We can only measure time by relying on some mechanical or physical processes that can be observed, and it is quite "possible that there is no uniform motion by which time may have an exact measure."[7]

However, the existence of absolute time is necessary for the real equality of any two time intervals or for the real simultaneity of any two events, or, as Newton says, "the flow of absolute time cannot be changed … [t]he duration or perseverance of the existence of things is the same, whether their motions are rapid or slow or null." [8] Newton reminds us that this fact has long been recognized in astronomy where the absolute or "true" time is distinguished from the practical way of measuring with the help of the so-called equation of time. Namely, Ptolemaic astronomers corrected the inequality of solar days by relying on the supposedly uniform sidereal day.

The need for immovable space is similar. We cannot make sense of motion and rest solely on the basis of the relations of bodies, so there must be some immovable space with a fixed structure of places according to which we determine whether some object is moving or is at rest. Again, absolute space is not empirically accessible, so for practical purposes we have to choose some "immovable" body relative to which places and motions of bodies are determined. Newton again warns us that it is entirely possible that there is no body that is truly at rest and that we will have to do with an approximation. For example, motion in the planetary system is measured in reference to the fixed stars.

Newton clearly thinks that the idea of absolute time is more intuitive than the idea of absolute space since he further supports the latter by comparing it to the former: "All things are placed in time with refer-

[2] Newton, The Principia, 408.
[3] Newton, The Principia, 408.
[4] Newton, The Principia, 408.
[5] Newton, The Principia, 409.
[6] Newton, The Principia, 414.
[7] Newton, The Principia, 410.
[8] Newton, The Principia, 410.

ence to order of succession and in space with reference to order of position."[9] And just as the order of the parts of time cannot change, so the order of the parts of place cannot change either. But exactly this conclusion follows from the relativity of space if places are determined in relation to bodies that occupy them. Moreover, the parts of space are not only places of things, but also of themselves, and it would be absurd if something could move out of itself. Consequently, there must be absolute places in reference to which absolute motions are determined.

In the rest of "Scholium" Newton tries to justify the proposed distinction between absolute and relative motion, or absolute and relative rest. He argues that the two are distinguished from each other by their "properties, causes, and effects."[10]

The first criterion are properties of motion and rest. Whether a certain body is truly at rest cannot be ascertained from its position relative to other local bodies since we do not know whether they are truly at rest or moving. The same can be said about true motion, the parts of a body keep their position within it, but are truly moving since the body itself is moving; a body keeps its position relative to the bodies in its vicinity, but it is truly moving together with them. To sum up, since the true state of rest or motion cannot be determined in reference to a body or to a relative, possibly moving, space, there must be an immovable space in reference to which absolute motion is determined.

The second criterion are their causes. True motion can be generated or changed only if force is impressed upon a body, while relative motion can be generated or changed irrespective of the applied forces. A force is applied to one body, but not to the other, thus producing a change in their relation. Consequently, the relative motion of both bodies changes, although only the former is subject to true change. Or, a force is applied to a body, but because the force is also applied to other bodies in the vicinity, the body is truly moving, but, relative to these other bodies, it is at rest.

The third criterion are their effects. The true circular motions cause "the forces of receding from the axis of circular motion."[11] Newton argues that the centrifugal endeavors of bodies, or its parts, to recede from the rotational axis are proportional to the quantity of true circular motion, while in purely relative circular motion they are null.[12] The claim that the true rotational motion cannot be just a relative motion is further supported with the bucket example.[13]

The bucket is hanging on a long rope. You turn it around until the rope is tightly twisted and then fill the bucket with water. At this point the bucket and the water are at rest. Next, you push the bucket into the opposite direction and let the rope unwind. The bucket begins to rotate, but at first the surface of the water stays flat. This means that the water is still at rest, but relative to the bucket, it is rotating. After a while, due to the friction of the rotating bucket, the water slowly begins to rotate as well, and its surface takes a concave shape, rising increasingly up the sides of the bucket with the accelerating speed until the water's rotation coincides with that of the bucket. At this point its motion relative to the bucket ceases, i.e. it is relatively at rest, while its true circular motion has reached its peak, as revealed in its endeavor to recede from the axis of motion.

The bucket example thus shows that the true rotational motion cannot be defined as its rotation relative to surrounding bodies. At the beginning, when the motion of the water relative to the bucket is the greatest, there are no centrifugal effects, as witnessed by the flat surface of the water – it is not truly rotating. At the end, when the water relative to the bucket is at rest, the centrifugal effects are the greatest, as demonstrated by its concave shape. This shows that the water is then truly rotating. And this rotational motion can also be determined from the magnitude of the centrifugal effects it causes: "The truly circular motion of each revolving body is unique, corresponding to a unique endeavor as its proper and sufficient effect."[14]

3. The proof of the existence of absolute rotation and absolute space

One interpretation of the bucket experiment is that Newton intended it as proof of the existence of absolute rotation and consequently, of absolute space. The experiment purportedly shows that the water does not rotate relative to the bucket, but does so truly, or absolutely. Absolute motion presupposes absolute space; therefore, the experiment also demonstrates the existence of absolute space.

This is the most common interpretation amongst physicists, as well as philosophers and historians of

[9] Newton, The Principia, 410.
[10] Newton, The Principia, 411.
[11] Newton, The Principia, 412.
[12] Newton, The Principia, 412.
[13] Newton, The Principia, 412–413.
[14] Newton, The Principia, 413.

science from Newton's time until the 1970s. Given that Newton himself admits that absolute space cannot be observed, it is not surprising to see many commentators conclude the experiment is a failure, and find Newton guilty of employing empirical means to prove metaphysical claims, thereby unjustifiably introducing speculation into physical science, which ought to attend only to observable and measurable phenomena.

Berkeley, for example, in his *De motu* (1721), insists that since we cannot perceive or even imagine space completely devoid of any bodies, absolute space is not real, and is actually "mere nothing."[15] Motion, Berkeley contends, can only be understood in connection with some determination or direction, but the latter cannot be understood if the moving body were the only body in the universe: "if we suppose the other bodies were annihilated and, for example, a globe were to exist alone, no motion could be conceived in it."[16] Similarly, we cannot conceive of the rotating bucket on its own in the universe. However, if we admit the existence of other bodies, then the motion of the water in the rotating bucket is not really circular, but rather "strangely compounded" of many motions, its components being, for example, the rotation of the earth around its axis and around the sun.[17]

The most influential was Ernst Mach's criticism in *The Science of Mechanics* (1883). First, he complains that "Newton has again acted contrary to his expressed intention only to investigate *actual facts*" since absolute space and absolute motion are "pure things of thought, pure mental constructs, that cannot be produced in experience."[18] Yet, the principles of mechanics result from our experimental knowledge, and extending them beyond the boundaries of experience is unwarranted. Specifically, in the bucket experiment, Newton is not justified in inferring that centrifugal effects are caused by absolute rotation from the fact that the rotation of water relative to the sides of the bucket does not produce any perceptible centrifugal forces. How does he know that they are not produced by its rotation relative to, say, "the mass of the earth and the other celestial bodies"? And what if we ask how the experiment would turn out if the walls of the bucket had been thickened until they were "several leagues thick." There is, however, no point in asking such a question, since we have this specific experiment before us and the task at hand is "to bring it into accord with the other facts known to us, and not with the arbitrary fictions of our imagination."[19]

Mach basically objects to drawing any conclusions that are not firmly based in experience and cannot be tested experimentally. The job of the experimenter is to show how the results of the experiment fit with other known facts and not to come up with some empirically unsubstantiated explanations. As Mach argues, absolute space and absolute motion certainly count as such and, accordingly, should be rejected. Many have been persuaded by Mach's argument, including Ian Hacking, Ernest Nagel, Hans Reichenbach, and Max Jammer.[20]

Similarly, Einstein sees the bucket experiment as a good illustration of an "inherent epistemological defect" of Newtonian physics. Causally inert space has no legitimate place in physics: "The law of causality has not the significance of a statement as to the world of experience, except when observable facts ultimately appear as causes and effects."[21] Indeed, even many of Newton's sympathizers do not bother much with the assumption of absolute space and motion, since Newtonian physics makes do with its empirical approximation. Or, for example, Carl Neumann replaces Newton's absolute space with an "unknown rigid body Alpha," which is at least a bit less abstract than absolute space. Its existence is as guaranteed as the existence of "luminiferous aether or of the electrical fluid" since "the possibility to refer all displacements to *one and the same object*" is essential to absolute motion.[22]

Objections can be shortly summed up in the following question: Why should we posit absolute time, space, and motion if we can construct the theory without them? This objection is, indeed, pressing, since observed phenomena may be neatly connected to relative, common, and apparent time, space, and motion, without introducing the concepts of absolute time, space, and motion.

[15] Berkeley, De motu, § 53. Relevant passages from Berkeley and others here mentioned authors can be found in a useful collection *The Concepts of Space and Time*, edited by Milič Čapek.
[16] Berkeley, De motu, § 58.
[17] Berkeley, De motu, § 62.
[18] Mach, The Science of Mechanics, 229.
[19] Mach, The Science of Mechanics, 232.
[20] For more information, see Laymon, Newton's Bucket.
[21] Einstein, "The Foundation of the General Theory of Relativity," 113.
[22] Neumann, "On the Necessity of the Absolute Frame Reference, 126.

4. The definition of absolute rotation within physical theory

Mach certainly thought that mechanics could be constructed solely with the concepts of relative time, space, and motion. For example, he concludes his analysis of the bucket experiment with the claim that centrifugal forces, i.e. the effects of rotation, depend on the relation of a body to all other masses in the universe and that any reference to absolute space is quite redundant. Einstein incorporated this idea into the general theory of relativity, maintaining that the choice of a frame of reference is arbitrary, thus proving that space-time and motion are inherently relative. Consequently, Newton's views on space and time seemed even more out-of-place than ever.

However, it has turned out that Einstein was wrong. Indeed, his general theory of relativity is not entirely relative after all. Its space-time structure is a function of the mass-energy distribution and assuredly is not "homogeneous and immovable." Nonetheless, it is definitely not a mere abstraction from relations among bodies. It is objective, and it shows itself in the phenomenon of motion. Thus, in Einstein's theory, as in Newton's, there are certain distinctions between states of motion that remain independent of the frame of reference one chooses and that arise from the structure of space-time itself. Consequently, it has become clear that the conceptions of time and space are not only a metaphysical supplement to physical theory, but an integral feature of a complete scientific account; they are not arbitrarily added to the theory, but rather, they follow from its laws. And since they are presuppositions of the physical theory they are influenced by the changes in our understanding of the physical world.

Starting from this deeper understanding of the role that the concepts of time and space play in a physical theory, Stein developed in "Newtonian Space-Time" (1967) a detailed analysis of Newton's mechanics and how it connects with his views on time and space. He persuasively argues that in his "Scholium" on time, space, and motion, Newton does not try to prove on the basis of physics that time and space are absolute, as the common interpretation would have it, but instead, seeks to determine their definitions within his physical theory in order to delineate both their empirical content and proper application.

And aAccording to this interpretation, the bucket experiment is not meant to prove absolute rotation, but to show that the concept is in fact grounded in the laws of motion entailed by the theory, and furthermore, that the distinction between absolute and relative rotation is indeed empirically justified. So, the proper answer to Mach's question – how do we know that the water in the bucket is rotating relative to absolute space and not to "the mass of the earth and the other celestial bodies"? – is to dismiss it as irrelevant. Newton actually claims that absolute rotation is defined, not proved, by centrifugal forces. Or, as DiSalle explains it, Newton "has defined a theoretical quantity, absolute rotation, by exhibiting how it is detected and measured by centrifugal effects."[23]

, since it is founded in his laws. We can always measure it by means of centrifugal forces, provided, of course, that the laws accurately describe the physical world. His definition of absolute time is also successful, since it clearly specifies the conditions in which two time intervals are equal even though we may never actually measure them precisely. This is, however, not the case with space, since the laws are not able to distinguish between absolute motion and rest. This is a well-known fact that Newton himself acknowledged, n which follows directly from the fact that in Newtonian physics force is measured by acceleration and not by velocity.[24]

An essential part of this interpretation lies in understanding how Newton's work connects with the work of his contemporaries and with the kind of physical questions they were grappling with at the time.

5. Newton in conversation with his contemporaries

It is common to interpret Newton's views on time, space, and motion in connection with Leibniz and against the background of the absolute vs. relational dispute that transpires, especially within the Leibniz–Clarke Correspondence (1715–16).[25] However, as Stein points out,[26] in the "Scholium," the main reference point is Descartes' "mechanical" philosophy from *Principia philosophiae* (*Principles of Philosophy*, 1644). It was the most influential physical theory at the time, and Newton wanted to

[23] DiSalle, "Newton's Philosophical Analysis of Space and Time," 45.

[24] For more information, see DiSalle, "Newton's Philosophical Analysis of Space and Time."

[25] See, Leibniz, Philosophical Papers and Letters.

[26] Stein, "Newtonian Space-Time." See also Laymon, "Newton's Bucket Experiment."

prove that, as opposed to his theory, it was unable to satisfactorily explain physical phenomena.

Descartes' theory is relativistic and his view on time and space is relational. According to him, the essence of material substance is its extension, i.e. its three-dimensional spatial extension, which basically means that material substance is identified with space. Additionally, since there can be no space without matter, the impossibility of vacuum follows. This leaves one with a difficult puzzle to solve, namely, how to explain motion; indeed, if the space that a body occupies is also its matter, then the body cannot change its place. Descartes tries to solve this problem by adopting the relational theory of motion and space: the motion of a body is its motion relative to other bodies. A body is thus at once involved in many different motions depending on which body is chosen as a reference point. However, there is one among them which is privileged, namely the true motion or, as Descartes calls it, "the motion in the strict sense." In *Principles of Philosophy* it is defined as "the transfer of one piece of matter, or one body, from the vicinity of the other bodies which are in immediate contact with it, and which are regarded as being at rest, to the vicinity of other bodies."[27] In short, true motion is a change of a body's position with respect to the bodies in its immediate vicinity.

Now, the bucket experiment can be read as a critique of Descartes' view with the help of a concrete example. The analysis clearly shows that his theory gives us wrong, even opposite, results from what we would expect based on our observations while conducting the experiment. When the rotation of the water according to its dynamical effects (i.e. endeavors to recede from the rotational axis) is the greatest, the water is, according to Descartes, at rest (relative to the bucket) and vice versa. Newton does not mention Descartes by name, but the whole example is structured in such a way that it clearly addresses Descartes' strict definition of motion and its failure to correctly capture this phenomenon. Newton thus proclaims: "true and absolute motion cannot be determined by means of change of position from the vicinity of bodies that are regarded as being at rest."[28]

Such a conclusion is further supported by the fact that Newton proceeds by applying the findings from the bucket example to the theory of planetary motions, which is a clear depiction of Descartes' vortex theory. According to this theory, a planet is at rest with respect to the contiguous tiny matter composing the solar vortex, which carries it around the sun, since the position of planet to contiguous matter remains the same in this case. However, this notion is clearly contradicted by the facts. Obviously, planets are truly moving: firstly, they change their relative positions with respect to each other, and secondly, they participate in the circular motion of the vortex–as its parts, they endeavor to recede from the rotational axis.[29]

The general difficulty is that Descartes treats the planetary system as a dynamical system composed of causally interrelated parts, while his definition of motion is kinematic, taking into account only changes concerning geometrical relations of bodies, and thus being too narrow for all kinds of motion that the vortex theory is meant to explain. Robert DiSalle sums this up helpfully by pointing out that Descartes' true definition of motion is not able to distinguish "between inertial motion and motion under the causal influence of force."[30]

A historically sensitive reading of "Scholium" and of the bucket experiment clearly reveals that Newton's purpose is not to derive metaphysical conclusions from physical theory and empirical observation. He sets out to explain the phenomenon of motion he himself and other 17th-century physicists, or natural philosophers as they were called, observed, by employing principles that involve suitably constructed concepts of motion, place, and space. These are not some metaphysical implants, but theoretical concepts with empirical content, which, indeed, must be revised if and when they are not able to participate in the satisfactory physical explanation of physical phenomena. In "Scholium" Newton analyses the relevant concepts used in the *Principia* and shows that the analogous concepts of the then prevalent physical theory by Descartes fails where his theory succeeds.

Newton's criticism can actually be applied to other physical theories from his time. Clearly, one cannot admit that forces affect the motion of bodies, but then, consider motion only in terms of changes in the spatial positions of bodies. The position is inconsistent, as Newton shows based on the example of Descartes' theory. However, one ought to show sympathy to their dilemma. Aristotelian deniers of the Earth's motion had a host of arguments to the effect that, if the Earth were really rotating or re-

[27] Descartes, Principles of Philosophy, 233.
[28] Newton, The Principia, 411.
[29] Newton, The Principia, 413.
[30] DiSalle, "Newton's Philosophical Analysis of Space and Time," 38.

volving, we should be able to observe certain phenomena attesting to this fact, t while there were none to be seen. Copernicans responded by claiming that there is no perceivable difference between the Earth rotating and the sphere of fixed stars being still, or the other way around. On the other hand, the collapse of Aristotelian celestial mechanics–the loud crash of the crystalline spheres–meant natural philosophers had to come up with a new explanation for the "circle-like" movements of planets around the sun and of the way they causally form an interconnected planetary system/world. This was the case, even more so, because they believed the "natural tendency" of planets is to travel in straight lines. Clearly, they reasoned, there must be something that causes planets to circle the sun. But what could this cause be? It was a difficult position.

Newton approached the problem analytically. The bucket experiment persuasively revealed the inconsistency in Descartes' relativistic definition of motion. He proceeded to show in which way the concepts of space and motion needed to be developed in order to support the dynamic model of the solar system/world that 17th-century natural philosophers envisioned. Therefore, I agree with Stein that Newton's analysis of time and space in "Scholium" is "a classic case of the analysis of the empirical content of a set of theoretical notions," and not an unwarranted use of empirical experiments in order to prove metaphysical theses.

6. Thought experiment or merely experiment?

In the literature on thought experiments the bucket experiment is regularly cited amongst the exemplary scientific thought experiments. But is it really a thought experiment at all? What Newton describes is an experiment that can be conducted by anyone, no special equipment being needed. In the text he even indicates – "as experience has shown me"[31] – that he himself performed the experiment, or was at least present when it was performed.

Why would so many philosophers working on thought experiments think that it is a thought experiment, even highlighting it as an exemplary one? Now, if the second interpretation of Newton's project is accepted, there is no reason for this at all. According to this interpretation, the job of the experiment is only to show how absolute rotation can be distinguished from relative rotation, and how it can be measured.

However, according to the first interpretation, Newton aims to prove the existence of absolute space – an unobservable, causally inert entity. In this case, clearly, the immediate results of the experiment do not suffice. They only present the first stage of the thought experiment: based on the described situation we conclude that the water does not rotate relative to the walls of the bucket. Then, in the second stage, we look for other possible explanations, only to realize that the water rotates relative to absolute space, thus proving the latter's existence.

And It is because of this speculative stage, in which Newton supposedly draws the conclusion about the existence absolute space, that commentators treat the bucket example as a thought experiment. The main question, of course, is why Newton disregards other possibilities – why is the water not rotating relative to the room in which the experiment is conducted, for example, or relative to Earth, or to great stellar masses? Why exactly is it absolute space that is chosen? Many critics devised thought experiments of their own: for example, Mach made the walls of the bucket many kilometers thick, Stallo came up with a lone star in an otherwise empty universe, and Einstein proposed a version of the experiment involving two rotating spheres.

To conclude, as a thought experiment, the bucket experiment is not very successful, and persuaded very few of the existence of absolute space. I believe I have presented a strong argument to support the claim that this, however, was never Newton's intention. Despite this fact, , the mistaken interpretation of Newton's project and its goals has resulted in many interesting thought experiments. It even helped Mach develop the idea that inertial effects depend on the relation of a body to all other masses in the universe, which, in turn, inspired Einstein to develop his general theory of relativity. In this sense, we can be grateful for the dynamic discussions on time, space, and motion that scientists and philosophers have participated in, and continue to participate in, even when their engagement has been founded, in part, on error and confusion.

References

Berkeley G (1976) Criticism of Newton: Excerpt from *De motu*, §§ 52–65. In *The Concepts of Space and Time: Their Structure and their Devel-*

[31] Newton, The Principia, 413.

opments, edited by Milič Čapek, 267–271. Dordrecht: Springer.

Descartes R (1985) *Principles of Philosophy*. In *The Philosophical Writings of Descartes, Volume 1*, translated by John Cottingham, Robert Stoothoff and Dugald Murdoch, 177–291. Cambridge: Cambridge University Press.

DiSalle R (2004) Newton's Philosophical Analysis of Space and Time. In *The Cambridge Companion to Newton*, edited by I. Bernard Cohen & George E. Smith, 33–56. Cambridge: Cambridge University Press.

Einstein A. (1952) The Foundation of the General Theory of Relativity." In A. Einstein, H. A. Lorentz, H. Minkowski and H. Weyl, *The Principle of Relativity*, 111–164. Translated by W. Perret and G. B. Jeffery. New York: Dover Publications.

Leibniz G W (1969) *Philosophical Papers and Letters, 2nd Edition*. Edited and translated by Leroy E. Loemker. Dordrecht: Kluwer Academic Publishers.

Laymon R (1978) Newton's Bucket Experiment. *Journal of the History of Philosophy* 16.4: 399–413.

Mach E (1919) *The Science of Mechanics: A Critical and Historical Account of its Development* (Fourth Edition). Translated by Thomas J. McCormack. Chicago & London: The Open Court Publishing Co.

Neumann C (1976) On the Necessity of the Absolute Frame of Reference. In *The Concepts of Space and Time: Their Structure and their Developments*, ed. Milič Čapek, 125–127. Dordrecht: Springer.

Newton I (1999) *The Principia: Mathematical Principles of Natural Philosophy*. Translated by I. Bernard Cohen and Anne Whitman. Berkeley: University of California Press.

Stein H (1970) Newtonian Space-Time. In *The Annus Mirabilis of Sir Isaac Newton 1666–1966*, ed. Robert Palter, 258–284. Cambridge, Mass. & London: The MIT Press.

BALKAN JOURNAL OF PHILOSOPHY. Vol. 11, Issue 2, 2019

LANGUAGE RELATIVITY

Serghey Gherdjikov

Sofia University, Faculty of Philosophy
Sofia, Bulgaria
e-mail: gherdjikov@gmail.com

Abstract:

We produce language forms via their *relations in coordinate systems*: languages. That is *virtual* language relativity. Languages are related to phenomena and work in the real life of communities. That is *real* language relativity. We use languages via symbolic behaviors, living in human communities. Relativism collapses at the level of successful exchange of experience between humans belonging to distant cultures. Relativism is a stance of not recognizing the *real relatedness* of all languages to one and the same human form and world. Absolutism (Universalism) is a stance of not recognizing *relativity as definiteness*, that is, the virtual interrelatedness of all languages. Languages are shaped by *human* life processes. We follow the path from "local languages," which are analogous to 'inertial systems', (this represents 'virtual relativity,' which is analogous to *special relativity* in physics) to *living people* talking about one shared sensual world (this represents 'real relativity').

Key words: relativity, language, world, real, virtual, translation, indetermination.

Introduction

Humans use speech in a *life-process of re-synthesis of the human form*, and this life-process is one and the same across cultures – that's why we *can translate natural meanings*.

Languages are shaped, ('infrluenced') by local, but still *equally universally human* life processes. From 'pure languages' (analogous to 'inertial systems') that make statements about the world and express a shared reality by means of translations (this represents virtual relativity,"' which is analogous to *special relativity* in physics) we move on to *living people* talking about one shared sensual world (this represents 'real relativity'). Human communities speak in forms that cohere with their local life-processes. But that 'relativism' collapses at the level of successful exchange of experience between humans belonging to distant cultures. Languages are mutually translatable. This is a point of view analogous to Einstein's general relativity, which describes states and movements beyond 'inertial systems'. This definiteness brings us beyond relativism.

A mother tongue is 'transparent,' or invisible, as one of the languages – i.e. there is an absence of being aware of the mother tongue with respect to how it refers to the world. We take our local grammar and categories as worldly entities. (The eye is invisible to itself.) Thus, "the optics" of the native language is also *undefined for its speakers*. The interference of language into the world remains unseen.

In any culture we have a local 'virtual sphere':, the space-time of artifacts and especially sign systems, or languages. These are shaped by local forms of human life, and represent a repository of our sign-forms. Franz Boas formulated this *relatedness* (not relativism) as a methodological rule: "An objective, strictly scientific inquiry can be made only if we succeed in entering into each culture on its own basis" (Boas, 1928, p. 205).

Anthropological relativism has a limitation (it appears to proscribe the notion of a universal 'human nature'. But anthropologists still use the notion/metaphor of 'cultural worlds'. This imposes a heavy conceptual conflict. Philosophers like Willard Quine, Thomas Kuhn, Paul Feyerabend, and Richard Rorty admit the relativism of languages, theories, facts, and specific vocabularies. The problem with this kind of relativism is the absence of any relatedness to everyday life; this cleavage between human life and human language means the relativism of this stance is unlimited, and at the same time, unex-

plained. To overcome relativism, we must *define relativities by conceiving of* local languages as *coordinate systems*, rooted in one and the same world. In doing so, we may replace the opposition of 'culture–nature' with the clearer opposition: *virtual–real*. Here, the virtual is the antithesis of the real–the sensual world.

All human languages as alternative *coordinate systems* that map onto (intersubjective) features of reality can be co-related, and translations can be good enough for successful meaning exchange. Members of different and even distant communities can recognize shared space/time forms and processes in acting together in the same environment. Science is not completely free of relativity, but observation and experiment as *non-language sources of meaning* are reliable enough. We humans do not live in distinct, discrete cultural *worlds*.

Seemingly, relativity provides no foothold for beginning to understand the connectedness of different cultural forms, but in fact *the foothold is precisely in the relatedness and uniqueness of these manifold forms.*

Relativity and Relativism of "Worlds"

The attitude according to which *world, thinking, and language have the same form*; that we speak what we think and what there is; that language is but a manageable device through which we express our thoughts about an independent external; reality, is rooted in our non-awareness of the specificity of a local language. For Plato and Aristotle, for Augustine and Thomas Aquinas, for Spinoza and Descartes, for Kant and Hegel, for Husserl and Heidegger, for Frege and for the young Wittgenstein language reflects thought and thought reflects reality in a genuine (veridical) way. This view is not just philosophical tradition, but also a disposition,. It is what the form *logos* consists in.

We need a new understanding of *relativity* beyond universalism and relativism, in order to describe a multitude of languages that function in one and the same world. Things spoken are identified in mutual relations; their meaning is dependent on one another and is established through signs pointing to phenomenal forms/processes. *This* without *that* is not *this*. I without you is not I. *West* without *East* is not *West*. *Man* without other living beings is not man. That is not 'determination of thought from language', it is *virtual relativity*.

The world contains languages; life contains speech. Thinking is far more undefined than language, and therefore, beyond the scope of this discussion. Hence, I will not delve into subjectivity here. I argue we overcome relativism *by relatedness to a common (concrete) world*, independent of any description (real relativity). This world is characterized by something non-relative and 'universal' (for humans): the organized flow of sensual forms in a life process. Definitions, spoken or written, bring language virtual relativeness. The notion of 'Life process' is not pure of description: without being brought into relation with the non-living, life is not 'life,"'but just 'existing'. So, "the organized flow of sensual forms in a life process" is, again, a description, opposed to the described flow of signaled and defined 'things'. So, pure 'real relatedness' is unspoken. We have here an option: some constructions in all languages share the same meaning, and they are clear for any human preceptor. Linguists call these "language universalities".

There is a (limited) diversity of different grand (binary) oppositions, such as: subject–object, mental–physical, culture–nature. These pairs of concepts are virtual, synthetic forms, not real entities – i.e. the pair "subject–object" corresponds to the relation: human body–environment. Here, let us experimentally eliminate these grand oppositions. Instead of them, I introduce a single opposition: "virtual–real". In one and the same speech community, divergence of meanings is small; a *native tongue* is a common frame of reference (*coordinate system*) that appears to fix meaning with limited discrepancy.

By contrast, comparison of distant languages shows great divergence of meanings, and it is very difficult to identify "real universalities" beneath these differences. Yet, we are, in fact, still able to recognize other people's experiences via translations and the shared experience of acting together in the same environment. The greatest cross-cultural conceptual differences we find that we (and our metaphysics) assume to be *universal* and *common to all mankind are as follows*: "God–word", "mind–matter", "subject–object", "I – world", "freedom–non-freedom". This is extremely important for understanding ontology.

Human communities are not just *carriers* of languages that purport to represent the world. They use speech and writing in a life-process of re-synthesis of populations of *Homo sapiens*. This claim is consistent with a Darwinian perspective, but also analogous to Einstein's general relativity in the following sense: pure "inertial systems" ("pure languages") do not exist; global space is 'non-

Euclidian. This view-point takes us beyond relativism.

If language relativity has no end and perceptions are dependent on languages, then "different language communities live in different worlds." That is *real relativism*, a position commonly attributed to E. Sapir and B. Whorf. In fact, this is not Whorf's position. He uses the expressions "objective world", and "objective realm," indicating the existence of an objective external world that is not contingent on human speech and perception (see Whorf, 1957, pp. 62, 63, 64). But he did write that cultures are "worlds" and language communities have their own perception of the world. Whorf, unfortunately, fails to clarify *how far languages shape perceptions*. Moreover, he adopts and quotes Sapir's motto: "The fact of the matter is that the 'real world' is to a large extent unconsciously built upon the language habits of the group." Whorf uses expressions like "*habitual thought worlds*", (Whorf, 1957, p. 147) "*American world*" (ibid., p. 173), and "*Western world*" (ibid., p. 152) in quotation marks; so, his position on the mind-world relation is not definite – that is, he does not present clear a distinction between thought, language, perceptions, and the external world. In other words, he does not *define relativities* here. The position of Sapir and Whorf– and, more generally, all linguists working within that tradition– *does not define the distinction between "culture world" and "external world"*. Hence, 'language relativity' is not distinguished from 'perceptual and thought relativity'.

The huge diversity of usages of the term "worlds" in culture studies is a symptom of this oversight. All such expressions should be taken as metaphors. Contemporary relativistic philosophy, which incorporates the vague conception of 'worlds' in various senses, has grown to be quite extensive (Quine, Kuhn, Rorty, Dummett, Bas van Fraasen, and others). This relativistic philosophy has elicited much criticism, T mainly from figures working in the cognitivist and "generative linguistics" traditions (see Pinker, 1994, 57–60). However, there is much evidence in support of certain aspects of relativistic philosophy, including experiments that support linguistic relativity (Slobin, 2003; Bailey, 2004; Mithun, 2004). In contemporary Linguistic Anthropology, moreover, we find many extensive investigations of "*cultural worlds"* (see Duranti, 2004).

As previously stated, human communities are not just *carriers* of languages representing the world. Humans use language as a part of broader projects that support the existence, growth, and evolution of their communities and cultures. Though these projects and languages differ from community to community, the way in which language reinscribes community norms, values, and goals appears common to diverse populations. The diversity of human cultures is thus consistent with the universality of 'human nature'.

To reiterate, 'real relativism' can be understood in the following way: If language *relativity has no limit* and sense perceptions are completely dependent on languages, then "different language communities live in different worlds".

We can find the most basic relativist ideas in the writings of Quine, Kuhn, Feyerabend, and Rorty. It would be impossible to cover the entirety of the debate that rages within the "post-positivist" perspective. But d three important arguments within that tradition stand out:

1. "Two groups see different things when they look from the same point in the same direction" (Kuhn).[1] This, however, *is not the case* (within and across cultures): when looking at stars, earth, trees, animals, humans, or biological cells, all groups see the same thing. They could refer to "things" differently, and hence, the act of seeing itself could be separated from the act of identifying what is perceived. But perception itself is universal.

2. "If observations are "nailed down", then scientific progress is impossible." (Feyerabend).[2] That *is not the case* in real science. "Nailing down" (or 'fixing') is necessary for scientific procedures of (objective) measurement.

3. "Conversion from, or to, another point of view is not a matter of inference from previously

[1] Thomas Kuhn, The *Structure of Scientific Revolutions*: "In a sense that I am unable to explicate further, the proponents of competing paradigms practice their trades in different worlds. Practicing in different worlds, the two groups of scientists see different things when they look from the same point in the same direction" (Kuhn, 1966, p. 150).

[2] Paul Feyerabend, *Against Method:* "Cultures change, they interact with other cultures and the indefiniteness resulting therefrom is reflected in their worlds. This is what makes intercultural understanding and scientific change possible: potentially every culture is all cultures. We can of course imagine a world where cultures are well defined and strictly separated and where scientific terms have finally been nailed down. In such a world only miracles or revelation could reform our cosmology" (Feyerabend, 1975/1993, p. 272).

shared premises."[3] That *is the case* in real conversions, but it is a mistake to assume logical inference always shapes facts (experience), while experience never shapes logical inference.

These three points argue relativity and indefiniteness are unlimited, but at the same time leave them *unexplained*. If two cultures (or paradigms) are "two worlds", *in which world do the proponents of these respective paradigms share* experiences? Insistence on a linguistic, formal, clear *definition* of a world that itself exists independently of any linguistic and formal definition is just nonsense, a blind projection of the attitude of *logos*. Perceptions do not "found" sentences, and sentences could not be *inferred* from phenomena – they are *designed to refer to them*. We see colors, but we speak *about* colors.

The problem with this kind of relativism is not a total absence of any reference to a particular phenomenal world, but the failure to define denotation in its relationship to sensual phenomena. Sensual phenomena are not formally linked to words, numbers, and other symbols. What, then, is the relationship between phenomena and signs that refer to them (language)? Relativism provides no clear answer to this question.

To overcome *relativism*, we must *define relativities*. Seemingly there is no foothold in relativity, but, in fact, the foothold resides in the *relatedness of different cultures and symbolic systems.* Anthropological relativism has a limit– a 'human nature' is postulated that may transcend observed differences among people. This provides a new understanding of the Absolutism–Relativism opposition: it is an opposition founded on non-discrimination between virtual and real relativity and relatedness.

Scientific observation is globally reliable as a source of knowledge that is not language-specific or culture-bound. Experiments and trans-cultural, 'non-grammatical' sign-systems like mathematics are maximally free from cultural and linguistic relativity – e.g. there is no 'Chinese physics' or 'British biology'. But the philosophy and theories (for example, sociological or anthropological theories) that may build on scientific discoveries may contain some culture-bound, local (*Western*) categories and concepts, which are not definite (with fixed, universally-agreed upon definitions) at the level of observation and experiment: (e.g. the concepts of matter, mind, thinking, and to some extent, natural laws, the objective world, observing subject).

We have here a clear distinction between real phenomena (that science deals in) and virtual strata of perception (when phenomena are interpolated into theory). So far, the theory of 'real relativity' has not been developed, i.e. a theory describing the curved space of global communication and description of the world while avoiding its reduction to the level of classical understanding. This 'general' theory overcomes the indeterminacy of the 'special,' or virtual theory, as I will show. It also overcomes the absolutism that dominates classical philosophy and aspects of theoretical physics, and that also characterizes contemporary relativism.

Virtual language relativity

Virtual language relativity is the definition of language constructions in relation to other language constructions/expressions in a specific language (*coordinate system).*

We may also speak of the '*Virtuality' of 'the thing'.* As Hilary Putnam put it: "'Objects' do not exist independently of conceptual schemes. *We* cut up the world into objects when we introduce one or another scheme of description." (Putnam, 1981, p. 52). Beginning with Eduard Sapir (1929), we can find the term "sociocultural worlds" in many variations of use, both metaphorical and literal. Whorf uses "nature" as a basis or ground for his discussion of worlds, stating, for instance:

> We dissect *nature* along lines laid down by our native languages. The categories and types that we isolate from the world of phenomena we do not find there because they stare every observer in the face; on the contrary, the world is presented in a kaleidoscopic flux of impressions which has to be organized by our minds – and this means largely by the linguistic systems in our minds. We cut nature up, organize it into concepts, and ascribe significances as we do, largely because we are parties to an agreement to organize it in this way – an agreement that holds throughout our speech community and is codified in the patterns of our language (Whorf, 1940, pp. 213–14).

[3] Richard Rorty: "…it is not that we live in different worlds than the Nazis or the Amazonians, but that conversion from or to their point of view, though possible, will not be a matter of inference from previously shared premises" (Rorty R (1991). "Solidarity, or objectivity?". *Objectivity, Relativism, and Truth*, p. 31)

It is a fact that our cultural experiences *organize* sensual forms, but it is not a fact that sensual *forms* are dependent on specific natural languages. There is not just a "flux", there is a *biological organization* of colors, sounds, smells, etc. Colors and their nuances are the same in different language communities with different color vocabularies (see Berlin and Kay, 1969). We have empirical science with exact instruments of observation and mathematical conceptualization, which *transcends local natural languages*, and there is not any necessary and insurmountable misunderstanding between English speaking and Chinese speaking scientists and readers of scientific articles. (Here, I won't discuss the topic of "incommensurability" in Philosophy of Science in any detailed way–again, see Berlin and Kay, 1969.)

It is true that we cannot know for certain what another person's subjective, qualitative experience of perceiving is like, independently of their verbal testimony and actions. Cognitive science observes (external) forms of symbolic and bodily behavior, rather than 'seeing"' and 'thinking' (necessarily private, internal experience) itself. Scientists observe individuals and groups belonging to different language communities, and if there is not 'incommensurability', or untranslatability, of these behaviors, then, it appears, there is some *global human form of sensual perception* and cognition . Furthermore, if we *can successfully act together* with individuals from alien language communities in a neutral environment, this indicates "we see and hear and otherwise experience" the same world. This is a fact on the level of *genetics* and ethology – there is a distinct genotype and behavior of the species *Homo sapiens*.

There is fierce criticism of language relativism–mostly from the point of view of cognitivism and generative linguistics–and it is supported by observations (see Pinker, 1994, p. 58). There is also strong support for relativism, and much evidence to defend linguistic *relativity* (see for instance Slobin, 2003, pp.157–192). We cannot deny the huge number of observations of Indian languages or Mandarin Chinese and Japanese, which affirm the *relativity of lexis, grammar, and ontology*.

It is reasonable to accept facts of universality (globality) and of unicality (locality), when they are reliable (and before they are subjected to the interpretive lens of any grand, totalizing theory). This controversial area can be clarified in light of the *distinction between virtual and real (language) relativity*. We do have global (universal) sensual, perceptual, and linguistic forms, and we do have local, culture-bound (unique) perceptual and linguistic forms; but it seems that we do not have concepts good enough to define them together and understand them as compatible with one another.

Real language relativity

The real (as opposed to the virtual) is just actually experienced phenomena. Here 'world' is the space/time in which the human life process flows. The real world is what we are born in, what we live, and what edie in.

The real world, however, is not identical to the '*Sensual world'*. The real world is not language, nor is it the 'sociocultural world.' Moreover, the "world in itself" is not in front of us (as 'objective reality'). Humans live in the environment, and the human body and environment are *experienced as a phenomenal world*. This position is a kind of "realistic phenomenology".

A color-form without sound, smell, taste, and touch melts into 'sensory form' blue without non-blue melts into bright; "brightness" without darkness melts into indefinite light. There are mutually related cells in the retina – three kinds for long, middle, and short length waves ("blue", "green" and "red") – and three kinds of bipolar cells for red–green, blue–yellow, and black–white in the chromatic P-strata in the visual cortex. Color receptors work together. The documented experience of "color illusions" shows that *color sensations are interrelated*. So, every color-quale is related to the others, and all of them to "frame of reference" *groups of color-receptors; the relationships between sensory forms in perception are thus like the grammar of a language.*

Speech is inseparable from breathing and from the movements of vocal chords, mouth, and lips. The information carried by speech points to a form/process or state/doing. Speech is a vital act and an intentional behavior, not just the interchange of information. Speech does not stand against an external world, but rather defines external phenomena. Grammatical form is adapted to communal life.

Wittgenstein, in his *Philosophical Investigations* (Wittgenstein, 1953), develops a philosophy of language, where language is life, opening the horizon for a philosophy of language *beyond logic and epistemology*, and beyond an ontology of the 'Universal'. In doing so, he extends his epistemology beyond an adherence to *logos*.

The purely logical, abstract conceptualization of language is replaced by a new understanding of language's diverse roles and dimensions in "*life*". "Speech act" replaces 'sentence' (e.g. language as "performatives"; see John Austin) as the primary bearer of meaning.[4] This revolution in the philosophy of language did not arise by chance, but was preceded by new insights in anthropology and linguistics.

"*Linguistic determinism".* Statements such as, "forms of thought determine forms of language" and "forms of language determine thought" are equally unreliable. These speculative claims are uninformative when checked empirically because (at the very least) both directions of dependency can be affirmed. That is why critics reasonably deny any one-way dependency of forms of thought on forms of language and vice versa (linguistic determinism). Language *organizes*, but does *not actually determine* thought and perception. Whorf embraces determinism not only in the direction speaking–thinking, but also in the direction mental–symbolic. According to Whorf,"My own studies suggest, to me, that language, for all its kingly role, is in some sense a superficial embroidery upon deeper processes of consciousness, which are necessary before any communication, signaling, or symbolism whatsoever can occur" (Whorf, 1956, p. 239). However, *thought–language* determinism (Frege, Wittgenstein in *Tractatus*) is a mostly speculative theory. So, it seems to be with contemporary cognitive theory of "mental language" in connection with "generative universal grammar" (Chomsky, Fodor, Pinker).

It is reasonable to develop an *analogy* of Special and General relativity (or of virtual and real relativity) within which a multitude of inter-related languages are related to human form, including language instinct and symbolic adaptive behavior. We never observe 'thinking' and 'thought' – we observe symbolic (verbal) and bodily acts. (Note that even private acts of introspection must be expressed in symbolic action that is informed by public processes of communication.) It appears more appropriate, then, *to relate language constructions to observable individual and collective behavior*, a choice that also allows us to sidestep wholly speculative discussions of language determinism' or 'universal grammar.'

Local language relativity

Real communication is virtually an informational exchange, and this exchange is a real inter-individual process of re-synthesis (organization and reproduction in the broadest sense) of groups of humans. Through real communication we share all kinds of experience: sense impressions, attitudes, emotions, and thoughts, and we are also able to coordinate our behaviors.

A real form/act is a moment in a life process. A local/momentary language construction/expression or bodily behavioral act is a moment in a local collective human life process. When moments acquire meaning,t they are"bent", or "distorted" when transmitted from one community to another.

Languages are shaped, ("curved") by local human life processes. From 'pure languages' (analogous to "inertial systems") making neutral statements about the world and transcending each other through translations (*virtual* relativity is here analogous to *special* relativity) we move on to *living people* communicating about their common world (real relativity, *general* relativity). People speak in forms that complement the forms of their behavior. A formulation in the opposite direction is: "Each language *shapes world*, describing it in its own form." But this proposition contradicts the facts: *successful exchange of experience between humans belonging to distant cultures is, in fact, possible.* All cultures are *forms of human life, and.* all human populations have natural languages. There is some kind of "language instinct" (Pinker) that is universal, despite the fact that we do not observe one "universal grammar."

"Language gravity". Languages are shaped by human life processes. That shaping is not a linear causal process; the shape a language acquires is not a product of determinism. *This suggests a process of "language inertia*." We experience that as the *stability of local languages*, which may we think of as analogous to the (relative) stability of the genotype of a population. Each local tongue *resists* translation into distant languages. Constructions in it are adapted to local forms of common living. People from distant cultures understand each other *with difficulty*; they may miss or even resist the unfamiliar, culture-bound messages that are not easily integrated into their existing worldview. Inertia (just as

[4] An illocutive act is the performance of an intentional linguistic act with certain conventional power. These concepts, no doubt, represent vital communication processes. Contemporary linguistics uses a general notion of 'constructions', which covers all linguistic forms of expression (see Goldberg, 2003).

in Einstein's General theory) thus shaping local speech as the "field" of meaning and experience.

Words are involved in the world. Willard Quine begins his analysis on reference (*The Roots of Reference*, 1974) by stating: "Bodies are not given in our sensations but are only inferred from them" (Quine, 1974, p. 1). This is recognized by Eduard Sapir as a "language behavior" and by Wittgenstein as a "form of life".

Words "involve themselves" in the world as they separate phenomena from one another by naming them.. This is so prevalent that, despite the incommensurability of speech and sense, we often can hardly distinguish between the name and the thing named. The senses distinguish things *without requiring the assistance of names*, and we learn to see before we learn to speak, but the borders between objects, and the definitions of things, are additionally drawn by local natural language and local common behavior. The *chair* is an artifact brought forth into the world by its *shape related to its function* ("meaning"). People who do not make chairs, do not sit in chairs, and have never seen a chair cannot perceive a chair as a chair. But they do perceive the shape and may eventually recognize its function.

Acts of speech and writing are "definitions" of collective life processes: *re-synthesis* of social attitudes, knowledge, customs, rules, activities. In the community there are fundamental vocal and written texts: ceremonial phrases; ritual written signs; prayers; instructions on agriculture, hunting, and animal husbandry; orders; rules for administrative life, legal proceedings, property certification, and so on.

Global language relativity

Real language relativity: language constructions are defined in relation to phenomenal forms/series and connected in the form of a life process (against the backdrop of world as the organized flow of phenomena) (see definition). A local/momentary language expression is a moment in a local collective human life process in the world.

Human life has the same (global) form as the re-synthesis of human form (see definition). With differences taken into account, languages are translatable, meanings are communicable, and the world a sphere of shared understanding for all people.

Sensual forms/events, approximately identical for human individuals across communities, are "gravity masses" that are denoted by words in manifold languages. f The same is true of biological forms of behavior: feeding, dressing, shelter, emotions, love, kinship, battles, hunting, exchange of legends, knowledge, etc., – we find words and constructions for all of them in all communities. Thus, no matter how distant, *languages converge towards a universal human form*.

There is a "classical domain" of a *homogeneous and isomorphic linguistic sphere* where statements and communications between speaking individuals and communities come into being. That domain is analogous to the "Newtonian" domain in mechanics, where we can neglect "curvature" of language according local human forms of social life.

In communication, linguistic understanding is attained by means of life meaning, in *real relativity*. Communication, in the sense of real relativity, corresponds to gravitational and accelerated movement in General relativity. Communication "curves information" by its very accomplishment: we understand others *by means of our own experience.* In fact, *this is no curvature, but a straight line of communication* in a "non-Euclidean" space.

Thanks to human form, meanings can be re-synthesed wherever humans live (with local "deformations"). Globality is a network of all life processes, especially of human cultures. Locality is a network of all life processes in a community, situated with respect to other communities. Real relativity allows us to define the world using the linguistic and conceptual tools provided by all sorts of languages and cultures. That's why it is a mistake to see cultures as distinct worlds. Analogously, "general relativity" in mechanics allows the formulation of laws of motion for all sorts of coordinate systems – inertial as well as non-inertial.

Hence, we may posit S the meaning of "culture and language universalities". First of all, these are sensual forms: colors, sounds, smells, etc. "If a language has only two colors – and all languages have at least two – they are always white and black; if a language has three colors, the one added is red", writes Donald Brown (Brown, 1991, pp. 13–14)". Brown lists many human universals:

> To begin with a few examples of human universals, those in the cultural realm include myths, legends, body adornment, daily routines, rules, concepts of luck and precedent, and the use and production of tools; in language there are grammar, phonemes, polysemy, metonymy, antonyms, and an inverse ratio between the frequency of use and the length of words; in the social realm there are a division of labor, social groups (including

> thinking of them as entities or agents), age grading, the family, kinship systems, ethnocentrism, play, exchange, cooperation, and reciprocity; in the behavioral realm there are aggression, gestures, gossip, and facial expressions; mentally there are emotions, dichotomous thinking, wariness around or fear of snakes, empathy, and psychological defense mechanisms." (Brown, 2004, p. 47)

Contemporary linguists continue to emphasize that there is a privileged entity in the realm of culture: language. "Observations about specific linguistic constructions have shaped our understanding of both particular languages and the nature of language itself" (Goldberg, 2006, p. 3). This view is not an extravagant one; rather it lies within a tradition of investigation that can be attached to the generative stream of contemporary linguistics, which springs from the works of Noam Chomsky (see, for example, Chomsky, 1957).

Josef Greenberg lists forty-five language universals in a form of "rules", for example: "Universal 1. In declarative sentences with nominal subject and object, the dominant order is almost always one in which the subject precedes the object" (Greenberg, 1963, p. 108).

These universals are not strict or absolute, i.e. they have exceptions:

> Universal 23. If in apposition the proper noun usually precedes the common noun, then the language is one in which the governing noun precedes its dependent genitive. With much better than chance frequency, if the common noun usually precedes the proper noun, the dependent genitive precedes its governing noun… However, outside of the sample there is at least one exception, Chinese, a prepositional language in which the relational expression precedes the noun (ibid.).

But if Chinese (Mandarin), the most commonly spoken language on earth, is an exception, then what can Greenberg mean by the term "universal"? As it turns out, "language universals" are difficult to formulate. The human brain contains cortical hierarchies for categories and for speech organization, so perhaps the surest universal here is that *Homo sapiens* share a genotype–yet, even this is not a strict universal: there are no two human individuals with the same genotype.

An ontology of the "Universal" is not adequate for capturing cultures or their languages. Lest we neglectWittgenstein's "families of language games", let me note that languages on earth are grouped into *families* (that is a fact*)*, and this kind of diversity is describable in terms of forms and relations.

Semiotic paradox: The sign "stands for" the object.

According to the classical semiotic definition, every sign "stands to somebody for something in some respect or capacity" (see Peirce, 1931, 2.228). J. Lyons U. Eco also relies on the formulation: "stands for." A sign cannot stand "for" or "in place of" anything except for another sign. The smoke in "that smoke means fire" does not stand for fire (natural meaning, Grice). The relation of 'standing for' is indefinite, fixed only in particular incidences.We can resolve this indefiniteness by replacing the "subject–object" opposition with the opposition: real–virtual. A *virtual* sign *defines a real* phenomenon, without causing it, or interacting with it. It is very similar to "function": a function of some structure, not necessarily present at a particular moment, is defined by it, and therefore,a virtual aspect of the structure. The word "red" does not really stand in for the quale *red*. The function of 'standing for' thus signals we have encountered intertwined life processes: looking and talking, which cannot stand in place of the other. The word "red" stays together with the word "green" as a pair of "colors" in our language (talking), and "colors" *virtually refer to the real* qualia of color vision (looking). The sign of a color is just a virtual "correlate" of unspoken quale. Semiosis, the product of these intertwined processes, is the defining of a sign together with other signs with respect to a quale, phenomenon, notion, act, operation.

Syntax–semantics–pragmatics

Syntax has its semantics: it is the real meaningfulness of constructions as language forms/series. The distinction between syntax and semantics/pragmatics is relative.

Grammar, virtually taken, is independent, but is semantically definable in a form of a world/life process: "somebody–does–something" (SVO). Grammatical categories have their semantic definitions, but do not come down to semantics. Grammar and semantics are not understandable outside the pragmatic context – it is the use of signs that, in the final account, confers on them meaning and references. Far from being engendered by grammar, behavior is

its very source. To take the 'universal grammar' as the basis of speech is to take the virtual for real. Grammatical categories have their semantic content via life forms and processes.

Grammar is local. Grammars alien to our own can use an adverb as subject, two subjects (with *-wa* and *-ga*) instead of 'subject–object' (Japanese) and a noun form as predicate (Chinese). In the Chinese sentence (translated literally as) "My notebook red" the subject is "my" and the predicate is "notebook red" and there is no verb at all. In Chinese the parts of speech are not strictly defined before the sentence is formed to provide a fixed syntax. Chinese linguists prefer to use a 'topic–comment' distinction instead of "subject–object," or "actor–action"; moreover, they can change their order, and all syntax is contextual.

As far as syntax (or semantics) shapes ontological entities, we have some "local ontology", indicating ontological relativity. Here Chinese does not demonstrate any fixed ontologically relevant syntactic structure. This, in part, explains the near complete absence of *speculative ontology*, or metaphysics in Chinese intellectual history.

Grammar traces out some aspects of 'local ontology' within which people belonging to a linguistic community 'order the world'. It supplies them with categories like 'noun– verb', 'subject–predicate', 'subject–copula–object' 'subject–verb–object'. Grammar has meaning as a network of forms of life processes. Those categories *are not a priori*, evident; they are not fixed and are not used in all languages in the world but are bound to certain cultural and historical contexts. Local "ontology" is *the form of descriptions* made in some natural language.

Translation and "indeterminacy"

Taken virtually, translation is a transfer of information from one language into another. If sign-systems are parallel to each other, or could be strictly coordinated, information may be transmitted from one system to another, e.g., one and the same number can be written in different numeral systems. In this process no information is lost.

But natural languages are not like that: they do not have parallel, strictly coordinated grammars. They are grouped into families, within which translations may be well defined; however, translations from one language family to another may *change information*. Nevertheless, such translations exist, and may be reliable enough. The fact that this is so cannot be explained by a "different worlds" theory. We have here a global field of local languages representing *one and the same world*. Translations between natural languages are not determined according to universal rules, but are adjusted thanks to *one world relatedness* (global real relativity).

A local/momentary language expression is a moment in a local collective *human life process in the world*. This is real relativity. As a real process, translation is the *transmitting of meaning*. Translation represents the transfer from one coordinate system (frame of reference) to another. And that transfer is not *determined* – it is "adaptation" via series of specific descriptive units. Some constructions get lost in translation insofar as they express forms and processes unique to the original speech community. However, communication between different language communities indicates mutual descriptions of different local forms (behaviors). This is possible, for example, in the experience of *cross-cultural living: participating* in a foreign community and sharing foreign forms of behavior, while still occupying the same sensual world. In this case, though people share the sensual world and engage in communication and shared undertakings, there is an unknown amount of indefiniteness (indeterminacy) with respect to how much is truly translated across cultures, and how much is lost in translation.

This relativity problem is described by both Boas and i Quine. Quine introduces "indeterminacy of translation" in his *Word and Object* (Quine, 1964), and later clarifies it in the article "On the indeterminacy of translation again" (Quine, 1987). Recognizing the relativity problem, Quine states, "An effect of recognizing this limitation is that the enterprise of translation is found to be involved in a certain systematic indeterminacy" (Quine, 1964, ix). He continues,"The firmer the direct links of a sentence with non-verbal stimulation, of course, the less that sentence can diverge from one to another from manual to manual… It is in this last form, as a principle of indeterminacy of translation, that I shall try to make the point plausible" (ibid., p. 27). "Our linguist would construct his manual of translation by conjectural extrapolation of such data, but the confirmations would be sparse (Quine, 1987, p. 6)

There is some *entropy* in translations: loss of information. Virtually we cannot recognize what is lost: there is no accounting for what has changed in the re-coding. But in reality we can compare patterns among messages and translations and overcome misunderstanding, talking at cross-purposes, distor-

tion, and annihilation of meanings. We may then transfer meanings via correct descriptions of (alien) life processes. Both the linguistic 'entropy' problem and its resolution are unexplainable from the points of view of absolutism and relativism. Absolutism turns out to be ethno-centrism.

So, cross-cultural experience, global communication, and especially investigations into languages and human communities, guide us to *deep rethinking and reformulations* in philosophy of language, ontology, epistemology, philosophy of science, and ethics.

Relativity of description

"Descript" is a word or symbol carrying the status of a category of a local description of the world (or artificial sign-system). There are many descriptive categories in one local language, which describe the world in a local, virtual, and real position. We may call a unit of such local description *descript*. There is a network of such descripts as *terms* in any particular descriptive system. By means of specific networks of descripts a local ontology, and a local life my be recognized. Sets of descripts represent local description in relation to others: *spirit, totem, deity*. Our classical philosophies use categories like *being, essence, idea*.

In a virtual and real global communication we have the challenge and chance to notice the *locality of* some western *leading concepts* (descripts) instead of their supposed *universality;* we have the opportunity to re-think and revise our "universal" conceptualization and ontology: *God–world, spirit–matter, mind–body, nature–culture, human–animal*, etc.

Science as a global enterprise leads beyond these oppositions. There is no empirical evidence supporting these descriptive divisions. Science uses concepts like *process–state, time–space, matter–energy, information–entropy, cell–organism, gene–expression, population–evolution*; they are organized in theories and work in experiments.

To take relativity of descriptions as "relativity of worlds" means to eliminate science as global form of knowledge and technology, and to miss the chance for clarifying *discrimination between experience and speculation*. Empirical science cannot exist at all if it adheres to ill-defined local descripts, born of local beliefs and habits, local religions and philosophies. Here we see local and global relativity of descriptions as systems of beliefs and knowledge of the human race. Science is not absolute knowledge, but it is *globally (universally) related* to *Homo sapiens*.

We transfer real meaning (phenomena, acts) in a pure virtual sphere. Words, statements and descriptions are projected from their real reference area (the *sensual world)*, into an imaginary (abstract, subjective, phenomenal) area: finite→infinite, accidental→necessary, being→nothingness. A concept referring to something at hand is projected into its polar opposite: a concept for which nothing at hand is present. In this virtual conceptualizing, we can go far beyond experience, imaging and creating *fictional entities–as* human communities do all the time. The newest fictional entity is, perhaps, *artificial superintelligence*, based on contemporary technology and rooted in very deep dispositions in our cultural history. That is a kind of *virtual projection*. We go from a real referred sign (word, symbol, notion) to a virtually referred one.

Virtual projection spontaneously passes beyond the limits of validity, due to our unawareness of the limits of experience and the limitlessness of "travel" in pure virtual space.

Denoting the denotation

A brilliant demonstration of linguistic indefiniteness is given many centuries ago in *Chuangtsu*: "Words are not just wind. Words have something to say. But if what they have to say is not fixed, then do they really say something? Or do they say nothing?" (*Chuang Tsu*, ch. 2).

Language is not *determined* by the world; it is not an *image* of the world. Words and qualia (elementary experiences) are incommensurable. Words just *point* to such experiences, and this 'pointing' is undefined d: *it is the defining itself*. The best ways of defining are rich in scientific *data*, recognized as a result of critical tests.

Linguistic indefiniteness in a positive perspective means linguistic variety, flexibility, plasticity, and reasonableness. If what was said were fixed definitely, it couldn't possibly be meaningful in real life. A given local grammar cannot be *drawn* from the phenomena – *phenomena don't have grammar*. Nor can it be drawn from the forms of life of the respective community. Instead, a grammar is defined through a process of gradual synthesis–a kind of natural *selection*. Grammar changes over time.

No sign system can denote *denoting itself – that is, relating signs to qualia, a* relation that seems to transcend scientific explanation. Semantics, represented in a meta-language, in which a phenomenon

is represented by a name together with the name itself (white snow – "white snow", Tarsky) is not a representing of the denotation. There is no metalanguage capable of voicing a "word pointing to the phenomenon". Pointing to a sensible form and uttering the name does no more to avoid this situation: pointing is itself a sign – signaling "that here". We have five or six senses, and each of them "creates" its own, specific qualia – colors, sounds, smells, tastes, touches, and internal sensations. They are integrated in our cerebral cortex, and subsequently captured by language, which links or fuses concepts and sensations through the process of mental imaging and thinking. Speaking is not parallel to sensing, as seeing is not parallel to hearing. These processes all display distinct and different qualities.

The sensible world is not congruent with description, but is exactly this cross-modal picture, or reality, which natural language is created for. It contains no internal grammar of its own, but it is precisely its sensual form that is the source of linguistic grammars in their various forms. In the sensible world there is no definiteness comparable to the definiteness of words, but it is this world that sets the horizon of words. There is no single, universal description for phenomena that crosses cultural and linguistic lines, but people from distant cultures may reach a mutual understanding of these phenomena through perception and investigation of natural phenomena.

There is no logically or grammatically definite world that is logically and temporally prior to language. But there is a real sensualworld shared by humans.

Nature–culture

Franz Boas formulated relativity as a methodological rule: "The scientific study of generalized social forms requires … that the investigator free themselves from all valuations based on our culture. An objective, strictly scientific inquiry can be made only if we succeed in entering into each culture on its own basis" (Boas, 1928, p. 205). This is known in anthropology as the 'principle of relativism'. The term relativism is misleading here: Boas indeed recognizes a common human nature unites human beings from different backgrounds; however, the failure to clearly outline this nature invites unlimited relativism.

The opposition of *nature–culture* is not definite. Is natural language "nature" or "culture"? There is no definite answer. Humans have a "language instinct" coded in our genes. We have organs of speaking and hearing. We produce and perceive sounds of speech. But we don't understand language as merely a kind of sound exchange. Language brings information; information about the real world is coded in the sign-system of a particular language, and translatable within other sign-systems. These systems are not created intentionally, but every word and rule is created by a community in a long, evolving, and unpremeditated process. Here again, it is unclear where we may locate the border between the natural and the artificial.

Such a distinction is clear enough in the opposition of *real–virtual*. The *Real* is the experienced phenomenal order, or experienced world; the 'world' being the space/time of human life processes/forms– the global set of human forms/processes. The *Virtual* is the created order, or artefact: a collection of signs, sign forms, languages, cultures; 'culture' being the space/time of artefacts/creations– the global set of local cultures.

Created form, or artefact, has an aspect free of real 'matter': it is sign, symbol, structure within a network of other signs, forms, structures. In reality, every form is a form of some phenomena. *Information* is carried by some physical or biological structure or substrata. Human language is coded in the brain, in diverse vocabularies, in various grammars, in indefinite descriptions.

One cannot live in a virtual world. One is not born, does not live, and does not die in the realm of language. But language allows us to communicate the most basic, elemental, bare truths of reality.

References

Bailey B (2004) Misunderstanding. In *A Companion to Linguistic Anthropology*, ed. A. Duranti, 395–413. Blackwell Publishing Ltd.

Berlin B, Kay P (1969) *Basic colour terms: their universality and evolution*. Berkeley: University of California Press

Boas F (1928) *Anthropology and Modern Life*. New York: W. W. Norton & Company, 1986.

Brown D (1991) *Human Universals*. New York: McGraw-Hill.

Brown D (2004). Human universals, human nature & human culture. *Daedalus*, 133(4): 7–54.

Chomsky N (1957) *Syntactic Structures*. The Hague, Paris: Mouton.

Chuang Tzu (1964) *Basic Writings*. Watson B, trans. New York: Columbia University Press.

Duranti A. (ed.) (2004) *A Companion to Linguistic Anthropology*. Wiley.

Feyerabend P (1975/93) *Against Method.* London, NY: Verso.

Goldberg A. (2003) Constructions: a new theoretical approach to language. *Trends in Cognitive Sciences*, 7(5): 219–224.

Goldberg A (2006) *Constructions at Work: The Nature of Generalization in Language*. New York: Oxford University Press.

Greenberg G (1963) Some Universals of Grammar with Particular Reference to the Order of Meaningful Elements. In *Universals of Language* ed. J. Greenberg, 73–113. London: MIT Press.

Kuhn T (1966) The *Structure of Scientific Revolutions*. Chicago: Univ. of Chicago Press.

Mithun M (2004) The Value of Linguistic Diversity: Viewing Other Worlds throughNorth American Indian Languages. In *A Companion to Linguistic Anthropology* ed. A. Duranti, 121–140. Blackwell Publishing Ltd.

Peirce C (1931–1935) *Collected Papers*. Boston: Harvard Un. Press.

Pinker, S. (1994). *The Language Instinct*. London: Penguin books.

Putnam H (1981) *Reason, Truth and History*. Cambridge: Cambridge University Press.

Quine W (1960) *Word and Object*. Boston: MIT Press.

Quine W (1974/1990). *The Roots of Reference.* Open Court Publishing Company.

Quine W (1987) On the Indeterminacy of Translation Again. *The Journal of Philosophy*, LXXXIV(1): 5–10.

Quine W (1968) Ontological Relativity. *The Journal of Philosophy* LXV: 185–212.

Rorty R (1990) Solidarity, or objectivity? *Objectivity, Relativism, and Truth*. Cambridge: Cambridge University Press.

Rorty R (1989) *Contingency, Irony, and Solidarity*. Cambridge, Boston: Press Syndicate.

Slobin D (2003) Language and thought online: Cognitive consequences of linguistic relativity. In *Language in mind: Advances in the investigation of language and thought,* eds. D. Gentner, S. Goldin-Meadow, 157–191. Cambridge, MA: MIT Press.

Wittgenstein L (1953/2000). *Philosophical Investigations*. Oxford: Blackwell Publishing.

BALKAN JOURNAL OF PHILOSOPHY. Vol. 11, Issue 2, 2019

MATHEMATICAL UNDERSTANDING AND "WHAT IF THINGS HAD BEEN DIFFERENT?" QUESTIONS

Andrei Mărăşoiu

University of Virginia, Corcoran Department of Philosophy
e-mail: aim3gd@virginia.edu

Abstract:
According to Grimm (2014), we only understand a phenomenon if we know what other phenomena it depends on, and we identify dependencies according to how we answer "What if things had been different?" questions. I argue that this view meets with mathematical counterexamples. For, in mathematics, things couldn't have been different. I consider three replies Grimm may make, and argue they do not succeed.

Key words: mathematical understanding, mathematics, philosophy of mathematics, mathematical understanding, phenomenon, conterexamples.

Introduction

On Grimm's (2014) view, presented in Section 1, we only understand a phenomenon if we know what that phenomenon depends on. We understand how things stand because we know how they could have differed. That is, we know what depends on what because we're able to answer questions of the form "What if things had been different?"

However, in Section 2, I mention two instances of mathematical understanding that, I argue, are counterexamples to Grimm's view of understanding. The first is understanding the conclusion of a simple arithmetical argument on the basis of its premises. The second is understanding an axiom of set theory, the axiom of infinity. I argue that understanding this axiom, and understanding the arithmetical truths in the argument given, are not mediated by our ability to answer "What if things had been different?" questions.

In Sections 3–5, respectively, I consider three replies Grimm may make, which interpret "What if things had been different?" questions differently: a purely modal approach, a grounding approach, and a modeling approach to answering such questions. I will argue that none of these approaches leads to re-interpreting the examples given in Section 2 so that they no longer undermine Grimm's view.

The Conclusion explores the wider epistemological consequences of this dialectical situation.

1. Understanding as knowledge of dependencies

In this section, I summarize Grimm's view of understanding.[1] He writes:

> Since understanding seems to arise from a grasp of all these different types of dependence, it might therefore be better [to] claim that understanding consists… of something like "knowledge of dependency relations" (Grimm 2014, p. 341).

Consider an example, borrowed from Carr (1961). Suppose you set out to understand the 1905 Revolution (a social unrest brutally repressed by Czarist authorities in St. Petersburg in 1905). The occurrence of the 1905 Revolution depended on sundry factors: macro-economical, social, mass-psychological. Unions were prohibited and peasants newly arrived in the city had no workplace protections. Previous protests had been met with violence

[1]This view is in the background of much contemporary theorizing of understanding, such as Zagzebski's (2001) virtue epistemology of understanding, and Gopnik's (1998) view of understanding as knowing the causal structure of the phenomenon understood. Closest to Grimm seems to be Greco's view, for whom: "[T]o have understanding is to have systematic knowledge of dependence relations. To understand a thing is to be able to (knowledgeably) locate it in a system of appropriate dependence relations" (Greco 2014, p. 286).

from riot police, and unrest was growing. Workers wanted acknowledgment of their problems. The newly created social stratum of industrial workers was disadvantaged in terms of income and social status. Our understanding of the 1905 Revolution is partly constituted by our knowledge of the variety of its causes, as well as their interrelations. Efficient causation is only one kind of relevant dependence. Other kinds of dependence are temporal or logical relations (e.g., the 1905 Revolution preceded the 1917 Revolution; and it instantiates the concept of revolution). The 1905 Revolution *depends on* each of these because the Revolution might have not occurred at all, or might have occurred differently than it did, had any of these factors been absent or modified. Understanding is a matter of degree: we know more or fewer factors the Revolution depended on; and we know, in more or less depth, how each of them influenced the Revolution's unfolding.

The view that understanding requires knowing what the phenomenon understood[2] depends on flows from other of Grimm's commitments. Grimm (2010, p. 337) argues that "understanding is the goal of explanation," so that at the end of explanatory inquiry the epistemic agent is able to explain why the phenomenon understood occurred. Moreover, Grimm (2014, p. 341) adheres to Kim's view that "explanations track dependence relations" (1994, p. 183).[3] And we identify what a phenomenon understood depends on by answering "What if things had been different?" questions:

> Following James Woodward (2003), one promising way to spell out the requisite notion of the kind of achievement that is needed—hence the kind of grasping that is involved in understanding—is in terms of having an ability to answer ''what-if-things-had-been-different?'' questions. To have an ability to answer questions of this sort, Woodward argues, is to be able to anticipate the sort of change that would result in the thing we want to explain (the explanandum) if the factors cited as explanatory (the explanans) were different in various ways. Brian Skyrms… expresses this same point in a more metaphorical way. According to Skyrms, grasping involves having an ability to anticipate that ''wiggling'' one variable will characteristically lead to a ''wiggling'' of another variable. (Grimm, 2006, pp. 532–533).

For instance, if fewer poor people had moved from the countryside to the city; or if the government had instructed the riot police to react with caution to occasional conflicts on the streets and in factories; or if better workplace protections had been secured; or if the workers had been allowed to unionize and start negotiating their position (etc.), then perhaps the 1905 Revolution might have been less bloody than it was, or might have occurred later than it did.

We can put Grimm's view of understanding as follows:

(1) If we understand a phenomenon, then we know how to explain why it occurs.
(2) If we know how to explain why a target phenomenon occurs, then we know what other phenomena the target phenomenon depends on.
(3) If we know what other phenomena the target phenomenon depends on, then we are able to answer "What if things had been different?" questions about those phenomena.
(4) So, if we understand a target phenomenon, then we are able to answer "What if things had been different?" questions about what the target phenomenon depends on.

Or, as Grimm puts it, understanding involves "an apprehension of how things stand in modal space" (2014, p. 334). The argument is valid. Presumptive counterexamples may be produced, which challenge the notion that understanding requires knowledge, or being able to explain what one understands.[4] But I am interested in what happens when everything goes well, epistemically speaking. So I will grant (1) for

[2]The word "phenomenon" is used broadly: any state of affairs, event, process unfolding, or property obtaining count as phenomena.

[3]As Kim elaborates: "[M]y claim will be that dependence relations of various kinds serve as objective correlates of explanations. Dependence, as I will use the notion here, is a relation between individual states and events; however, it can also relate facts, properties, regularities between events, and even entities. We speak of the "causal dependence" of one event or state on another; that is one type of dependence, obviously of central importance. Another dependence relation, orthogonal to causal dependence and equally central to our scheme of things, is mereological dependence: ...the properties of a whole, or the fact that the whole instantiates a certain property, may depend on the properties had by its parts" (Kim, 1994, p. 183).

[4]For discussion surrounding whether understanding a phenomenon requires being able to explain it, see Lipton (2009), Khalifa (2012) and Strevens (2013).

the sake of argument. Similarly, one may advance alternative conceptions of explanation, or resist any general view about scientific explanation; but I will grant (2) for the sake of argument.

I will argue in the next section that instances of mathematical understanding provide counterexamples to Grimm's view[5] because, in mathematical cases, one *cannot* go about answering the question "Why is that the case?" by answering the question "What if things had been different?" For, in mathematical cases, things *couldn't* have been different. And considering patent impossibilities seems explanatorily idle, rather than conducive to understanding.

As I granted (1) and (2) of Grimm's view, and since, in the next section, I offer two counterexamples to (4), I reject (3): the kind of understanding we have of phenomena we can explain doesn't always require that we are able to answer "What if things had been different?" questions. In Sections 3–5, respectively, I consider three replies Grimm may make, intended to re-interpret the instances of mathematical understanding I mention so as to ensure they do not conflict with Grimm's view. I will argue that the replies fail to achieve that end.

2. Two simple cases of mathematical understanding

In this section, I argue that two instances of mathematical understanding[6] are counterexamples to Grimm's view.

Initially, Grimm (2006) had advanced a conception of understanding as "knowledge of causes," according to which we only understand that which we know the causes of. However, this conception involved efficient causation alone, and couldn't be generalized to cover mathematical understanding. To remedy that, Grimm (2014) advocates a theory of understanding based on dependencies in general, rather than on efficient causation alone:

> given how closely our notion of causation is tied to pushing and pulling – to exerting causal force – a perhaps more attractive strategy would be to demote the notion of causation from its central role and instead to appeal more generally to the notion of dependence (Grimm, 2014, p. 341).

There are a number of problems with this passage that I bracket.[7] However, I will argue that the resulting view faces counterexamples.

For the first counterexample, consider the understanding provided by the following argument:

(a) Every strictly positive natural number has a predecessor.
3 is a strictly positive natural number.
So, 3 has a predecessor.

The argument is valid. The premises are true, so is the conclusion. Moreover, the premises jointly *explain* the conclusion. One way of *understanding* the conclusion – the instance of understanding I'll discuss – is based on the premises. I will argue that Grimm's account of the understanding we have of the conclusion on the basis of the premises in (a) is false.

On Grimm's view, an answer to the question "Why are things so?" that provides an understanding of the state of affairs explained is obtained from an answer to the question "What if things have been different?" Given that the premises in (a) jointly provide an understanding of the conclusion in (a), that understanding has to be obtained *via* (b) below, since (b) answers the "What if things had been different?" question with respect to (a):

(b) If 3 had not been a strictly positive natural number, or if not every strictly positive natural number had a predecessor, then 3 might have not had a predecessor.

Statement (b) is a mathematical counterfactual. How to interpret mathematical counterfactuals is too large an issue to settle here.[8] It's clear that, when we un-

[5]Grimm's extending the "knowledge of causes" view to mathematical understanding contrasts with Woodward's (2003, p. 3) explicit disavowal of applying his view of scientific explanation to mathematical cases.

[6]The discussion to follow doesn't presuppose that mathematical understanding is a unified category of epistemic evaluation. There are a variety of mathematical projects worth pursuing, and they largely differ in their characteristic epistemic features. See Avigad (2008) and Hafner and Mancosu (2005) for different kinds of explanation in mathematics, and, presumably, different kinds of attendant understanding.

[7]To consider only one issue: Our ordinary concept of causation *isn't* push-pull. We think demand may generate supply, Congress can keep the President in check, and a word can hurt our feelings; none of these are push-pull, and all are causal according to common sense.

[8]I submit it would be mistaken to think all countermathematicals are trivially true given falsity of antecedents. Consider: "If one [says] 'nothing sensible can be said about how things would be different if the axiom of choice were false,' it seems wrong…: if the axiom of choice were false, the cardinals wouldn't be linearly ordered, the Banach-Tarski theorem would fail and so

derstand why 3 has a predecessor on the basis of the facts that every strictly positive natural number has a predecessor, and 3 is such a number, what we are entertaining is (a), not (b). We have a grasp of the relevant dependencies (3 is a strictly positive natural number, and hence shares their properties) without any appeal to counterfactuals like (b).

Here is a second example: the set-theoretic axiom of infinity, according to which there are infinitely many things. This claim is necessarily true if true at all, for there doesn't seem to be any room to ask what the existence of infinitely many things might be *contingent* on.[9] Suppose, for the sake of argument, that the axiom of infinity is true. It follows that no possible worlds exist where it is false. It's hard to see how the mere inexistence of possible worlds falsifying the axiom would *explain* the axiom or help us understand it.[10]

I sketched two instances of mathematical understanding: the understanding afforded, *via* a simple arithmetical argument, of the conclusion by the premises; and the understanding we have of an axiom of set theory. These instances of understanding are importantly different. The first is within the grasp of grade school children; the second is more foundational. The first is arithmetical, the second is set-theoretical. The first is proof-based; the second is an axiom. If, as I contend, both are counterexamples to Grimm's view, this shows that the problems with that view are quite general.

In Sections 3–5, respectively, I'll consider three replies Grimm might make, in which the instances of mathematical understanding I just mentioned aren't counterexamples to his view. I will argue that these replies don't succeed in explaining away the counterexamples. As Grimm's view of understanding was meant to apply to all cases of scientific understanding – mathematics included – it should be abandoned, given the compelling counterexamples it faces.

3. A modal approach to answering "What if things had been different?" questions

In this section, I argue against a reply that Grimm could make countering the notion that the instances of mathematical understanding mentioned in Section 2 are counterexamples to his view. The reply is that we *do* have an answer to "What if things had been different?" questions. Namely, that they couldn't have been different. This answer is satisfactory with respect to the examples given in Section 2 because knowing how things *must* stand in the mathematical realm leads to an understanding of why things actually are that way.

I believe this reply misinterprets "What if things had been different?" questions. Consider a mundane example. In preparation for surgery, a patient is given an anesthetic. The dosage has to be right if the operation is to proceed seamlessly. We may ask: what if things had been different? Had the anesthetic dose been smaller, the patient might have woken up on the operating table. Had the anesthetic dose been much larger, the patient could have had a heart attack. By intervening on the anesthetic dose, we see effects in the patient's health condition. We know why the patient was sound asleep during surgery but woke up fine afterwards: it was partly because he had been administered the right dose of anesthetic. Intervening on causes often alters their effects, so answering the question "What if things had been different?" in this context leads to the discovery of causal relations.

By varying the anesthetic dose, we vary its effects. In contrast, in mathematics, there's nothing to vary, so long as we agree with Grimm's (2014, p. 334) claim that knowledge of mathematical truths is *a priori*. So "What if things had been different?" questions clearly must play a different role in understanding mathematical matters than in understanding others, e.g., how anesthetics work.[11]

forth" (Field, 1989, pp. 237–238). Field is right. To appreciate whether the conditional is true, we have to grasp a bit of set theory and topology. That isn't trivial, regardless of whether the Axiom of Choice is true or false.

[9]In the text, I only need the claim that the axiom of infinity is necessarily true if true at all. There are traditional ways to deny this claim (such as Mill's suggestion that arithmetical truths are inductive generalizations generated by counting pebbles), but they seem to me to lack any plausibility.

[10]In keeping with (1), one may wonder in what sense we may explain the axiom of infinity when we understand it. Maddy (2011, pp. 113–137) addresses the issue, considering a form of explanation specific to foundational axioms, which she terms "elucidation." How elucidation may differ from other forms of understanding is beyond the scope of this text.

[11]We should distinguish Grimm's view from a related one. "One way in which our wonderment about a phenomenon can be relieved is through a demonstration that it is necessary, that it could not be otherwise" (Glymour, 1980, p. 32). Sometimes truly understanding why something *actually* obtains requires knowing that it obtains *necessarily*. This is surely true sometimes, and may well apply often in mathematical matters. But it seems to obviate an appeal to the "What if things had been different?" question, rather than require it. Glymour is clear that this

Once we admit that mathematical statements are necessarily true if true at all, then answering questions of the form "What if things had been different?" is trivialized. Compare "If every natural number has a successor, then 1+1 exists" with "If infinite sets exist, then 1+1 exists." If both are necessarily true, then a purely modal approach will be unable to ascertain that the first expresses a genuine dependence whereas the second does not. Assuming that we do understand that 3 has a predecessor because all strictly positive natural numbers do, then either that understanding is not given *via* an explanation such as that in (a), *contra* (1), or that explanation doesn't refer to dependencies between arithmetical facts, *contra* (2). If a purely modal approach to "What if things had been different?" questions in arithmetic trivializes their answers, given that banalities explain nothing, it is hard to see how a purely modal approach could underwrite our understanding of arithmetic.

Moreover, if answers to "What if things had been different?" in mathematics are trivialized, then finding them is no epistemic achievement. Even though we may readily produce such answers, it is not thanks to the exercise of our cognitive abilities.. Grimm endorses (1), by which we only understand what we know how to explain. It follows that an element of *cognitive achievement* corresponds to coming to understand something new, and what is achieved is learning what the phenomena understood depend on. This shows that, if (3) is assumed, (1) is undermined.

One might reply that counterfactuals like (b) can be interpreted so that they are non-trivial, by considering not only metaphysically possible worlds, but also metaphysically impossible worlds in which mathematical truths could possibly be false (Nolan, 2014). Perhaps such a reply succeeds in making the attribution of a truth value to (b) non-trivial. But it doesn't alter the fact that (b) isn't required in order to understand the conclusion of (a) – that 3 has a predecessor – based on its premises: that all strictly positive natural numbers have predecessors, and 3 is such a number. And it is hard to see how an approach that considers metaphysically impossible worlds would *illuminate*, or enhance our understanding of the axiom of infinity.

In the next two sections, I explore whether two other ways of interpreting "What if things had been different?" questions are conducive to explaining and understanding phenomena inquired into. I will argue that this is not always so in the mathematical realm.

kind of understanding by knowing-it-is-necessary only holds *sometimes*.

4. Grounding to the rescue?

Grimm mentions grounding (2014, p. 341),[12] suggesting the possibility of elaborating a reply to the claim that Grimm's view faces counterexamples relying on a grounding-based metaphysics. In this section, I argue that appeal to grounding is ill-suited to re-interpreting the instances of mathematical understanding mentioned in Section 2.

The idea of appealing to grounding in order to fine-grain a modal approach to the dependencies grasped in understanding a phenomenon seems to be the following. To know what a phenomenon is caused by (*lato sensu)*, or depends on, when we understand that phenomenon, is to know the grounds of that phenomenon. The idea that explaining why the phenomenon understood occurs is accounted for by answering the question "What if things had been different?" takes on a new guise: vary the grounds, and you will have thereby varied the grounded. By building a profile of how the grounds of a phenomenon vary, one can explain why the phenomenon occurs as it does.

Wilson argues that a general notion of grounding isn't theoretically helpful in illuminating what metaphysical dependencies are:

> The problem here is not just that claims of Grounding (failure of Grounding) leave open some interesting questions; it is that such claims admit of such underdetermination — about whether the dependent goings-on exist, are reducible or rather distinct from the base goings-on, are efficacious, and so on—that

[12]In fact, Grimm mentions Aristotelian formal causation, as well. He regards the choice between appeal to grounding and appeal to formal causation as alternative accounts of metaphysical dependence in general as terminological (a diagnosis I don't share). Grimm states that "Aristotle's notion of causation was more expansive, along the lines developed here," (2014, p. 341) referring to Aristotelian causes, in their formal, material, efficient, and final varieties. What formal causes are is quite obscure, but that problem can be side-stepped because these varieties of causes easily generalize to the plurality of dependencies. Mancosu (1996) chronicles how geometers gradually gave up talking about formal causes as they began articulating a notion of demonstrative proof anticipating those current nowadays. This is why I focus on grounding as a way to articulate Grimm's view of how understanding involves knowledge of what grounds what.

> even basic assessment of claims of metaphysical dependence, or associated views, cannot proceed by reference to Grounding alone. As such, investigations into metaphysical dependence cannot avoid appealing to the specific 'small-g' grounding relations (again: type or token identity, functional realization, the classical mereological part–whole relation, the causal composition relation, the set membership relation, the proper subset relation, the determinable–determinate relation, etc.) that are capable of answering these crucially basic questions about the existential, ontological, metaphysical, and causal status of metaphysically dependent goings-on (Wilson 2014, p. 540).

Wilson's point is the following. The theoretical construct of Grounding (capitalized to distinguish it from specific dependence relations) is supposed to capture metaphysical dependence in general. But this theoretical construct can't by itself help answer specific questions: "whether the dependent goings-on exist, are reducible or rather distinct from the base goings-on, are efficacious, and so on." These questions, however, need answering. This is why one should look for another way to describe various metaphysical dependencies; and Wilson argues that part-whole, logical, causal relations etc. provide satisfactory answers to the questions that need answering. I endorse Wilson's criticism, and think it is the advocate of Grounding who has the burden of proof to show that Grounding can provide an illuminating account of metaphysical dependence.

How might a "small-g" grounding version of Grimm's view address the counterexamples in Section 2? Take the arithmetical counterexample first. The fact that every strictly positive natural number has a predecessor, and the fact that 3 is a strictly positive natural number, jointly *ground* the fact that 3 has a predecessor. This is an improvement over a purely modal approach, for now there are part-whole and logical relations that can be invoked as relations of metaphysical dependence going beyond whether the mathematical facts in question hold with metaphysical necessity or not. Because 3 is *a part of* the strictly positive natural numbers, the fact that every strictly positive natural number has a property (having a predecessor) *logically implies* that 3 has that property too. These logical and part-whole relations can do justice to the understanding-giving and explanatory asymmetry of the premises and conclusion of (a), improving over a purely modal elaboration of Grimm's view.

However, a grounding elaboration of Grimm's view has trouble accounting for (b). This is because (b) is an answer to "What if things had been different?", whereas grounding relates to metaphysical necessity, and it *doesn't* seem to make sense to ask what it might have been for metaphysically necessary facts (that every strictly positive natural number has a predecessor, that 3 is such a number, etc.) to be different from what they are. We lack an answer to the question of what it is to *vary* mathematical grounds. By the same token, we can't vary facts (taken to be metaphysically necessary) such as the existence of infinitely many things.[13] So a grounding-based reply in support of Grimm's view doesn't succeed in addressing the understanding we have of the axiom of infinity either.

I have considered an interpretation of "What if things had been different?" questions on which answering them indicates what grounds what, for varying the grounds would vary the grounded. And I have argued that, while this improves over a purely modal approach to "What if things had been different?" questions, it doesn't rescue Grimm's view from counterexamples.

5. Different models

Grimm (2016) argues that "mental models" play an important role in the cognitive activity that underwrites understanding. This raises the question of whether a modeling approach to understanding[14] could rescue Grimm's view from counterexamples. The guiding thought would be that, when we ask how mathematical truths could have differed, we contemplate alternative models we use. My reply is that, while we may at times contemplate a variety of models, that isn't always required for understanding the phenomena theorized.

[13]One might propose that the existence of infinitely many things is grounded in the continuity of spacetime. However, that spacetime is continuous (if it is) seems to be a claim of less modal strength than what the axiom of infinity requires, for it is a claim in the foundations of physics, not of set theory.

[14]Models, here, are conceived broadly, so as to include both representations internal to a thinker's mind – her memory of concepts or schemata – and external representations, models understood as cultural artifacts to be found in textbooks or laboratories. I depart from Grimm's reference to "mental models" in order to consider a view as close to Grimm's as possible that still remains noncommittal on the cognitive activity that underwrites understanding, for such psychological speculation would raise questions of plausibility on its own.

This text doesn't discuss a modeling approach to understanding *per se*, but only inasmuch as it offers a new interpretation of "What if things had been different?" questions asked in order to identify what a phenomenon depends on, so as to ultimately explain and understand that phenomenon.[15] A modeling approach wouldn't eliminate the role that "What if things had been different?" questions play. Such questions are important in the natural and social sciences, where counterfactual scenarios are a good guide to testing. Rather, we answer "What if things had been different?" questions *differently*, by considering how our models could differ in how they represent the same target phenomenon, even when that phenomenon obtains necessarily.[16] The word "model" has an astonishing variety of uses.[17] However, in order to speak to the counterexamples in Section 2, the view has to presumably consider models in the model-theoretical sense.

In the case of arithmetical understanding mentioned in Section 2, answering "What if things had been different?" questions has to meet two tasks. First, that we can make sense of the mathematical counterfactual (b). Second, that it is somehow *via* interpreting this counterfactual that we gain the understanding of the fact that 3 has a predecessor in (a). I believe that the first task is met but that the second is not.

Here are two models relative to which we may make sense of (b), or non-trivially assign it a truth value. First, consider a model of first-order predicate logic that has only two elements. This model would *not* also be a model of first-order Peano arithmetic. In this model, (b) is true because its antecedent is false. But the justification is not due to any metaphysical impossibility. Rather, the justification has to mention the *model* in which we interpret (b). Second, consider a model of two elements, but this time as a model of the positive integral domain $\mathbb{Z}_1$. On that model, for instance, $(4)_{\mathbb{Z}_1}=(2)_{\mathbb{Z}_1}=(0)_{\mathbb{Z}_1}$ and $(5)_{\mathbb{Z}_1}=(3)_{\mathbb{Z}_1}=(1)_{\mathbb{Z}_1}$. Mimicking the successor function for natural numbers, in $\mathbb{Z}_1$ the successor of 0 will be 1, the successor of 1 will be 0, and so on. "3" is interpreted as $(3)_{\mathbb{Z}_1}$, and we can ascertain that it does have a predecessor. Again, (b) will come out true because its antecedent is false. But, again, the falsity of the antecedent of (b) and the truth of (b) are established partly with reference to the model in which we interpret (b).

I'll now argue that showing (b) is coherent, and can be interpreted non-trivially, doesn't *also* show that the understanding (a) delivers of why 3 has a predecessor has to be mediated by (b).

We might think as follows. We can test the validity of the argument in (a) by considering models, trying to falsify the conclusion and see if the premises are thereby falsified or not. Since the argument in (a) is valid, no such model will be forthcoming. And it is open to say, with Grimm, that we might better *appreciate* the validity of the argument in (a) by considering models. In considering such models, *volens nolens* we consider models that would give (b) the same truth value, albeit for different reasons.

[15]We should distinguish the view considered in this section from Saatsi's (2018) view. Saatsi writes that "genuine explanations are underwritten by explanatory dependencies in the world. This is the basic factivity requirement of the counterfactual-dependence account. Explanatory understanding, in turn, can be construed as an agent's ability to make correct counterfactual what-if inferences. [And] what matters for explanatory progress is that understanding-providing theories and models *de facto* latch onto reality in appropriate ways, so as to satisfy explanations' basic factivity requirement" (Saatsi 2018, pp. 14, 2). What Saatsi seems to be suggesting is that we explain phenomena by drawing inferences from the models we have of those phenomena. And models represent, not just what there is, but also how the target systems would behave in different conditions. This differs from the view attributed to Grimm here, on which "What if things had been different?" questions consider, not varying phenomena given a fixed model, but varying models of presumptively the same target phenomenon.

[16]On a modeling approach to "What if things had been different?" questions, it doesn't matter that the antecedent of (b) may not pick out any metaphysically possible worlds. For we can ask: "What if things had been different – logically, set-theoretically, arithmetically, etc.?" And these questions would be neutral on metaphysical matters, showing that the third way of interpreting "What if things had been different?" questions is superior to the first two ways of interpreting them, sketched in Sections 3–4.

[17]After giving several quotes from science textbooks illustrating different uses of the word "model," Suppes writes: "The first of these quotations is taken from a book on mathematical logic, the next two from books on physics, the following three from works on the social sciences, and the last one from an article on mathematical statistics…One of the more prominent senses of the word missing in the above quotations is the very common use in physics and engineering of "model" to mean an actual physical model as, for example, in the phrases "model airplane" and "model ship" … It would, I think, be too much to claim that the word "model" is being used in exactly the same sense in all [these contexts]" (Suppes, 1960, p. 3). In arguing that "the meaning of the concept of model is the same in mathematics and the empirical sciences" (p. 4), Suppes is aware that he is going against first appearances.

Whether we put the matter in terms of the verbalized subjunctive conditional (b) is irrelevant. What matters is that we entertain such models in testing if (a) is valid.

There is some plausibility to this line of thought, but it surely cannot hold in all cases, and we have no special reason for thinking it holds in the case of (a)–(b). The key point to notice is that we may have a purely syntactic understanding of why 3 has a predecessor. We grasp that all strictly positive natural numbers have predecessors, and that 3 is a strictly positive natural number; we grasp *modus ponens* and universal instantiation, and can thereby deduce the conclusion in (a), that 3 has a predecessor. We need appeal to no models in doing so.

The reply generalizes: we may have a purely *proof-theoretic* understanding of a conclusion without appealing to modeling. To be sure, valid arguments are such that, in any model where the premises are true, so is the conclusion. But mention of models is indirect, and it is unclear why understanding the conclusion of a proof *could not* be achieved without reference to models.

I am not advocating a proof-theoretical approach to mathematical understanding. Intuitively, the proof-theoretical approach and the model-theoretical approach are complementary, and precisely which of them (and in what contexts) better serves understanding should not be judged before the fact. But this is precisely what (3) does in implying that we can only understand the conclusion in (a) *via* (b).

In this section, I argued that a modeling approach to "What if things had been different?" questions cannot rescue Grimm's view from counterexamples.[18]

Conclusion

On Grimm's (2014) view, we only understand a phenomenon if we know what its occurrence depends on. On this view, we know what depends on what, and are enabled to explain target phenomena, by answering "What if things had been different?" questions. However, I developed two counterexamples to Grimm's view of understanding, and argued against three replies Grimm may make.

The fact that Grimm's view meets with these counterexamples may suggest that different instances of understanding may vary significantly in their epistemic profile according to *what* is being understood: bits of mathematics, the occurrence of revolutions, or how the dosage of an anesthetic contributes to the success of an ensuing surgery. The variety of possible interpretations of "What if things had been different?" questions, and how we may answer them, reflects this. Whether an informative general, topic-neutral, epistemology of understanding does justice to the epistemic particularities of different instances of understanding is still an open question.

This dialectical situation is important because Grimm's view considered understanding in full generality, whereas the counterexamples I mentioned are instances of mathematical understanding. The details of Grimm's view aside, it is worth emphasizing that the question of what mathematical dependencies ultimately consist in is left open, even though the fruitfulness of a purely metaphysical analysis of such dependencies may be undermined.

For a simple instance, think of "1+1=2".[19] This

[18]To simplify things, the text focuses on the first counterexample to Grimm's view in Section 2. As for the second counterexample in Section 2, we might think that two images support the axiom of infinity. Either infinitely many points exist in spacetime (so infinitely many concrete individuals). Or we may build the iterative hierarchy of sets, which is as close as we may come to intuitiveness without running into inconsistency. (Field, 1980 and, respectively, Boolos, 1971 are the classical sources for these two images of sets.) Asking how things might have been different amounts to asking whether we may contemplate models that retain plausibility despite departing from these two images of what sets are. Finitism, for instance (even if *per impossibile*), is a live theoretical option. The point is not to definitively decide any issues with respect to infinity, but only to point to how basic theoretical and modeling choices inform the ways in which we understand the axiom of infinity. In reply, notice that calling the claim that there are infinitely many spacetime points, or the iterative theory of sets "images" plays no essential role. We might as well *deduce* the axiom of infinity from such claims and theories. Suppose we wrote the model theory for the Zermelo-Fränkel set theory in a meta-language of our choice, so that the "intuitive" iterative conception of sets is simply a theory in one's chosen meta-language, all of whose models are also models of the object-language Zermelo-Fränkel set theory being interpreted. All this is to say that no obvious benefit follows from appeal to "images," as opposed to just more theory.

[19]Alternatively, we may revert to the two counterexamples I have offered against Grimm's view. As for understanding that every strictly positive natural number has a predecessor, what is understood obviously depends on understanding what Peano's axioms express, together with logical operations and what the converse of a relation is (here, the converse of a successor-relation, viz., the associated predecessor-relation). As for the understanding the axiom of infinity, and the role that finite models may play in grasping the relations between set theory and arithmetic, see Krynicki and Zdanowski (2005). Since their dis-

states an elementary fact, on which much of mathematics depends. In its turn, this fact depends on what identity and addition amount to, as well as on what the nature of numbers is. However, what might we mean in saying that some of these facts "depend on" one another? We might say that 1 is a *part of* 2. If numbers were sets, we might say that 1 is *an element of* 2. If we adjoined two collinear unities, we would thereby *construct* a segment that is two units long. We may build a *deductive proof* of "1+1=2" from Peano's axioms and the usual logical rules and axioms. Number-structures that include a position for 2 must include *substructures* that include a position for 1. And so on. We may cast such elementary dependencies in terms of deduction, construction, parthood, membership, etc. Which of these applies for developing an r understanding of arithmetical facts, and in what contexts, is, I believe, still an open question.[20] And whether these would count as distinct dependencies, or whether they merely *represent* in different ways what amounts to the same dependence of 2 on 1 is another question I leave open.

What *isn't* left open here is a methodological choice going forward. When we consider "What if things had been different?" questions as they apply to mathematics, or when we consider what mathematical dependencies may amount to, we should try to spell out the metaphysical or metamathematical views available, and discuss their specific details. I have here sketched Grimm's view, and argued it doesn't afford a notion of mathematical dependence correlative to "What if things had been different?" questions that may illuminate what our understanding of mathematical matters consists in.

Acknowledgments

Previous drafts have benefited from insightful comments and questions from Brie Gertler, Jim Cargile and Paul Humphreys (all UVa), and the audience at the 2018 British Society for the Philosophy of Science meeting. I am also grateful to an anonymous reviewer. Financial support was provided by the University of Virginia through a Dissertation Completion Fellowship.

cussion is more mathematically sophisticated, I chose to end with the much simpler "1+1=2" to illustrate questions surrounding what mathematical dependence consists in. I am indebted to an anonymous reviewer whose suggestions prompted me to rethink the conclusion of this paper.

[20]For an elaboration of, and examples in support of, the view that understanding may depend, in mathematical contexts, on which specific problem we are in the process of solving, cf. Avigad (2008).

References

Avigad J (2008)Understanding Proofs. In *The Philosophy of Mathematical Practice* ed. P. Mancosu, 317–351. Oxford: Oxford University Press.

Boolos G (1971)The Iterative Conception of Set. *Journal of Philosophy* 68: 215–231.

Carr E H (1961) Causation in History. In *What is History?,* 113–143. New York: Random House.

Field H (1980/2016) My Strategy for Nominalizing Physics, and its Advantages. In *Science without Numbers,* 42–47. New York: Oxford University Press.

Field H (1989) Realism, Mathematics, and Modality. In *Realism, Mathematics, and Modality,* 227–281. New York: Blackwell.

Glymour C (1980) Explanations, Tests, Unity and Necessity. *Noûs* 14: 31–50.

Gopnik A (1998) Explanation as Orgasm. *Minds and Machines* 8: 101–118.

Greco J (2014) Episteme: Knowledge and Understanding. In *Virtues and Their Vices* eds K. Timpe,C. A. Boyd, 285–301. Oxford: Oxford University Press.

Grimm St R (2006) Is Understanding a Species of Knowledge? *British Journal for the Philosophy of Science* 57: 515–535.

Grimm St R (2010) The goal of explanation. *Studies in the History and Philosophy of Science* 41: 337–344.

Grimm St R (2014) Understanding as Knowledge of Causes. In A. *Virtue Epistemology Naturalized* ed. Fairweather, 329–345. Springer.

Grimm St R (2016) Understanding and Transparency. In *Explaining Understanding* eds St. R. Grimm, Chr. Baumberger, S. Ammon, 212–229. New York: Routledge.

Hafner J, Mancosu P (2005) The Varieties of Mathematical Explanation. In *Visualization, Explanation and Reasoning Styles in Mathematics* ed. P. Mancosu, 215–250. Dordrecht: Springer.

Khalifa K (2012) Inaugurating Understanding or Repackaging Explanation? *Philosophy of Science* 79: 15–37.

Kim J (1994) Explanatory knowledge and metaphysical dependence. *Philosophical Issues* 5: 51–69.

Krynicki M, Zdanowski K (2005) Theories of Arithmetics in Finite Models. *Journal of Symbolic Logic* 70: 1–28.

Lipton P (2009) Understanding without Expla-

nation. In *Scientific Understanding: Philosophical Perspectives*, eds H.W. De Regt, S. Leonelli, K.Eigner, 43–63. University of Pittsburgh Press.

Maddy P (2011) *Defending the Axioms.* Oxford: Oxford University Press.

Mancosu P (1996) Philosophy of Mathematics and Mathematical Practice in the Early Seventeenth Century. In: *Philosophy of Mathematics and Mathematical Practice in the Seventeenth Century*, 8–33. Oxford: Oxford University Press.

Nolan D (2014) Hyperintensional Metaphysics. *Philosophical Studies* 171: 149–160.

Saatsi J (2018) Realism and Explanatory Perspectives. In *Understanding Perspectivism: Scientific Challenges and Methodological Prospects*, eds M. Massimi and C.D. McCoy. London: Routledge.

Strevens M (2013) No understanding without explanation. *Studies in History and Philosophy of Science* 44: 510–515.

Suppes P (1960) A comparison of the meaning and uses of models in mathematics and the empirical sciences. *Synthese* 12: 287–301.

Wilson J (2014) No work for a theory of Grounding. *Inquiry* 57: 535–579.

Woodward J (2003) Introduction and Preview. In: *Making Things Happen: A Theory of Causal Explanation*, 3–24. Oxford: Oxford University Press.

Zagzebski L (2001) Recovering Understanding. In *Knowledge, Truth, and Duty: Essays on Epistemic Justification, Responsibility, and Virtue*, ed. M. Steup, 235–249. Oxford: Oxford University Press.

BALKAN JOURNAL OF PHILOSOPHY. Vol. 11, Issue 2, 2019

NOTES ON THE EXPRESSIVE STATUS OF PHOTOGRAPHY

Costin Popescu

University of Bucharest
Bucharest, Romania
e-mail: bhmors.popescu@yahoo.com

Abstract:

With a history spanning almost two centuries, during which technological advancements on the one hand, and the fluidity of social life on the other, have constantly posed new challenges, photography keeps redefining its place among visual artifacts, and its functions among self-regulatory societal mechanisms.

As a widespread practice, photography has generated analyses and commentaries from many theorists of various fields, as well as from more than a few working photographers. Arguments and judgments on the status of photography with regard to its expressive possibilities, adhere to vastly different and often divergent points of view, and most importantly, raise considerable difficulties that prevent these discussions from relying on any methodological coherence.

The present text presents some of these arguments and judgments. It aims to provide grounds for more orderly future debates on the artistic quality of photography and especially on the methods of investigating photographic artifacts.

Key words: isotopy, visual field/visual universe, resolution, framing, realism/formative tendency.

Introduction

In 1839, the French state purchased the patent for Daguerre's invention and made it available for general use. In 1941, William Henry Fox Talbot developed the calotype. Since then, due to constant improvements in photographic techniques, interest in photography has increased exponentially.

Debates on the status of photography are virtually as old as the practice itself. They engage photographers, art historians, aestheticians, sociologists, psychologists, semioticians. While the status of photography is probably the most debated topic (Cartier Bresson, 1952; Bourdieu, 1965; Berger, 1968; Sontag, 1977; Barrett, 1990; Maynard, 1997; Bate, 2016 …); it has generated another line of inquiry, namely the artistic quality of the photograph (Sontag, 1977; Barthes, 1980; Clarke, 1997; Bate, 2016…). In the realm of the arts and humanities, equally nuanced discussions address the issue of the methodologies employed in analyzing the visual artifacts themselves (Barthes, 1980; Floch, 1985, 1986; Kress, Van Leeuwen, 1996; Alexandrescu, 2008…).

Commentaries on the status of photography and its artisticity frequently compare painting and photography. To certain commentators, contrasting photography with painting serves to distance the former from the latter. Comparing and contrasting them almost never stresses what photography can do well (with the most notable exception: scientific photography). Even when painting is not mentioned in discussions of photography, the latter's potentialities and achievements are insufficiently noted: "photography transforms reality into a tautology. When Cartier Bresson goes to China, he shows that there are people in China, and that they are Chinese" (Sontag, 1977, p. 111).

If we are to discuss them comparatively, I find it more profitable to investigate how painting and photography derive different expressive effects from the *characteristics of visual perception*. For instance, one might experiment with the expressive means of the visual by employing different, equally vibrant color tones within the same figure (*La femme au chapeau*, 1905, Matisse) – where the different parts of the woman's face are equally bright, as proves the desaturation of the colors, but their different hues are striking (the portrait looks much more common once the painting is desaturated); portraits can be constructed by combining photographs together, so that

every part of the image is made out of several photos with a similar degree of luminance (Robert Silver creates such *photomosaics*).

I shall go on to discuss the characteristics of photography (always a topical matter, due to the ongoing diversification of the process of taking and exhibiting photographs), based on the idea that the shared common ground – the vision – of the above-mentioned realms of producing visual artifacts can reveal relevant similarities and differences.

*

Edward Weston has noted two qualities of photography:

First there is the amazing precision of definition, especially in the recording of fine detail; and second, there is the unbroken sequence of infinitely subtle gradations from black to white.(in Trachtenberg (ed.), 1980, p.172).

If, as some state, a veil obscures and hides the world from our understanding, and photography may remove it, (Sontag, 1977, p.120) and if photographers, as Edward Weston believes, can show others "what their own unseeing eyes had missed," this is probably due to the high resolution of photographic images. For the photograph surpasses human perception. Captured on the daguerreotype plate, the pavement and the dew droplets (as they appeared in a photograph – C.P.) demonstrated, through the astonishing clarity of previously unacknowledged objects, the limited nature of human perception (Krauss, 2013, p.81).

As early as 1839, an American journalist noted admiringly that examined through a magnifying glass, photographs evinced a precision of detail that could never be achieved by the naked eye (Kracauer, in Trachtenberg (ed.), 1980, p.246). Is it a paradox, then, that the possibility of revealing *objectness* (Bate, 2016, p.142) in such detail should lead to the need to go beyond it?[1] It is well known that biological perception is selective, but it has rarely been stressed, as Maynard notes (1997, 2p.07), that we are aware of this fact. Our perceptual experience greatly exceeds our perceptual knowledge; hence the fear that we are constantly missing something– that there are things we should see, but do not. Photography, however, allows us to hopefully compensate for this failure. Obtaining and contemplating such a clear image can give its author or beholder the sense of possessing a power, a miraculous energy that must entail a commensurate revelation. Such energy cannot be wasted; high resolutions invite photographers/beholders toward their objects, beyond their objects – to go deeper and deeper and discover in them forever more (*the greater mystery*, as Weston puts it). Isn't high resolution a solution for our reconcilement, however mediocre, with the limitations of our (visual) abilities? Is it too far-fetched to compare a high-resolution photographic image to the image of an object seen through a magnifying glass? Both terms of this comparison contain details inaccessible to the naked eye at the moment when the photo is taken, or the magnifying glass is put aside. If the photograph is said to lend importance to the object being photographed, this is even truer in the case of high resolution images. We suddenly become better-sighted. The precision of the photograph frees the object from the perceived banality we normally associate with it, and teases us by suggesting we have failed to see everything that was there to see (obviously, the value of a photograph is not simply given by its 'amazing precision of definition'). It is not unimportant that originally such clarity in rendering the texture of objects could raise fears; in the early days of photography, John Ruskin (1819-1900) declared that photographs reveal "too much", and thus drive us away from aesthetic beauty: "exaggerating the powers of sight necessarily deprives us of the best pleasures of sight" (Bate, 2016,p. 120).

This clarity of images encourages the pursuit of realism. We should accept this pursuit as very old. It is more than a mere anecdote that Zeuxis of Heraclea, a famous painter of Antiquity who owed his celebrity, as we learn from David Hockney – himself a painter – to his astonishing naturalism (Hockney, Gayford, 2017, p. 70). He asked Parrhasios of Ephesus, his equally realistic fellow painter, to remove the veil hiding his canvas, only to find that it actually was the *trompe-l'œil* representation of a curtain. Reproducing fragments of reality with such accuracy demonstrates the power of controlling them. Thus, photography successfully achieves man's old desideratum of securing a position of power (of course, in the meanwhile, man has been redefining positions of power).

Other commentaries on the specific qualities of photography, mention framing as a definitory trait:

[1] As Weston states: "I don't want the play of sunlight to excite the fancy, nor the mystery of gloom to invoke the imagination (…), I want the greater mystery of things revealed more clearly than the eyes see" (Newhall, 1975, p.40).

> As a finite object, the camera cuts out a portion of an infinitely larger setting (…) Once the photograph has been cut out, the rest is removed. The implicit presence of the rest of the world, and its explicit removal are as essential aspects in the practice of photography as what is explicitly shown (after Stanley Cavell, in Krauss, 2013, pp.158-159)

Terry Barrett provides a more thorough look into the matter:

> Photographs, by nature, are always swatches cut from seamless reality. They are segments, shot from close or afar, with a wide or narrow angle of view. By use of the viewfinder, photographers include or exclude. In a sense, all photographs are literally 'out of context.' They are out of a spatial context, and they are also out of a temporal flow. They are one instant stopped in time. To understand and appreciate photographs, it is sometimes very useful to imaginatively put them back into their *original contexts*, to see what the photographer has done to make a picture, to study what was included, and how, and to imagine what was excluded and why. To consider their temporal element, we can try to see them as if they were one frame from a feature length film. We can imaginatively consider what was physically available to the photographer at the time the exposure was made (Barrett, 1996. p.89).

The impressionists are most likely the first authors of pictures that seem to unfold beyond the frame. Manet's *Un bar aux Folies Bergère* lets us see in the mirror what the painter wanted to show from behind himself, including the man on the right who is talking to the woman in the centre. In Monet's *Plage de Trouville*, the female figure on the right of the viewer has her left shoulder cut out. Monet also painted the *Femme à l'ombrelle tournée vers la droite* and the *Femme à l'ombrelle tournée vers la gauche*; these women direct their gaze beyond the lateral sides of the canvas, prompting us to imagine *what* they are looking at, and *how* the space only partially shown in the painting continues to unfold. Could such instances be tokens of the influence of photography? Enough correlations between impressionism and photographic techniques have been signaled, to prove they are not merely coincidental.[2] And the matter of framing includes many aspects that deserve more attention than the amount attached to it by the commentaries of Cavell and Barrett.

Of course, in many cases the photographer restricts, or 'obturates' the visual field, confining as much as possible the beholder's attention from looking outside the image for something to alter the significance already constructed of this field (Fig. 1). Here, the author's concern is to fill the field with something meaningful enough for the "pact for significance" concluded with the beholder to work – but, as we have seen, high resolution paves the way for this pact; different degrees of focal clarity, the arrangement of shapes, etc. contribute in their turn; Berger adds to this the influence of the *general idea* (Berger, Mohr, 2016, p.126).

Fig. 1

Here are two of the above-mentioned aspects: a) placing in the visual field some figures or silhouettes gazing outside it (not including the unfocused gaze of such persons who are portrayed and who for various reasons, stare vacantly), towards something which becomes important for the viewer, precisely because of the gaze of the respective character; b) introducing in the visual field some traces or indications of the presence, beyond the visual field, of persons or things that enhance its significance.

[2] Here is an example: In the 19th century, the painting was defined as a constructed whole. For a painting to be recognized as such, it needed a unifying vision, which would structure a set of objects and place them in a particular relationship. The works of Daubigny disobeyed this rule. What kept Daubigny from transposing nature in paintings complying with the conventional norms of unity? The only answer that seems to lend some coherence to the situation is: the fact that he had encountered and experienced photography.(cf Krauss, 2013, pp. 80-81). Daubigny is a precursor of impressionism.

In *Kamaitachi no 8* (1965), Eikoh Hosoe takes the picture of a fence and has a man perched above it, on the highest tree branch. The crouching man is gazing into the distance. His tension is revealed by his crouching position, by his closeness to the right side of the picture (his gaze has no room to unfold), by the black clouds (the fence is just a set of odd bars unable to stop anything from passing, criss-crossing the white sky between the field and clouds), by the field itself where the fence stands and, if it extends further, prevents the perched man from seeing the person he expects to arrive. What is there to be seen? Should it be seen at all, or shouldn't it? Do we ask such questions with regard to a painting? How frequently does painting resort to this technique? Might photography reveal and make visible to us what painting does not?

Fig. 2

Christopher Michel is the author of a photograph where an elderly woman looks from the street into a shop window (Fig. 2). The beauty of it lies precisely in the way in which the space and the viewer's gaze are distributed within and without the picture. The old woman is standing outside the shop, in the street, looking in; she is also within the photo, while the object of her gaze – the shop – remains outside it. All of a sudden, the inscription on the door, which we can read backwards in the photo, loses its plainness: *No dogs allowed*. Does that old woman really have a dog that she wishes to bring inside, into the warmth? Is the author hinting at a dehumanization of the old woman (who is afraid to enter)? Is she entering without her dog? Is she alone, simply looking for someone, and so the significance of the text on the door is completely unrelated to the significance of her presence in front of it? We may reproach photography for such ambiguity of meaning; T but may we not regard photographs as visual chronicles of our own lives, which are often rather confused themselves? There are several layers of meaning: one of the interplay between inside and outside (shop /street), another between the visual field and the visual universe, and finally another between the elderly woman and the inscription on the shop's door. We may want to smile, remembering how often life presents us with such uncertainties (most of the time, we simply dismiss them as 'business as usual').

Many photos of sports competitions present the shadow cast by athletes. If an athlete is seen making a great effort, with his body twisted in order to make the necessary action, then his shadow in the picture requires us to reconstruct the twisted body; might we never have seen a body in such a movement? How closely does the reconstructed expression of the body resemble its images generated by my visual memory?

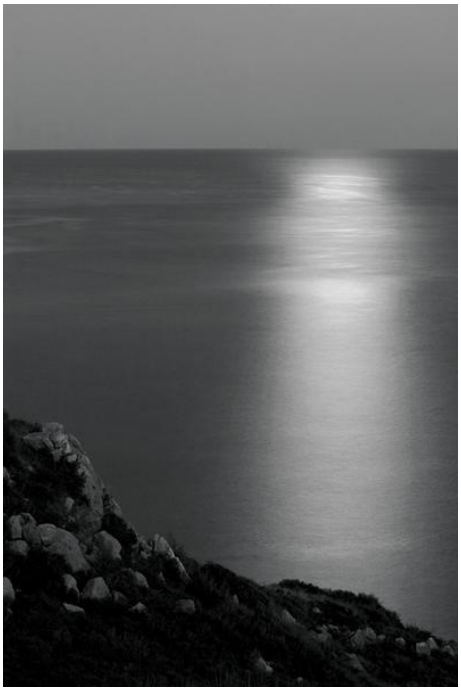


Fig. 3

Horia Măruşcă has taken a picture of a Greek landscape where the moon is missing (Fig. 3). Its light is perpendicular to the horizon line, spreading very simple lines across the zigzagged contours of a rock. Everyone knows what the moon looks like; however, in this particular photo, everyone would like to see it again: right here, right now, above this

shore. The contrast between the darkness of the night and the shining strip of light across the great surface creates an urgent desire to actually see the moon. And, indeed, , we almost see it.

Thus, photographs are able to create a major distinction: between the *visual field* (what the photographer has placed in the picture) and the *visual universe* (what is outside the visual field, but can be guessed or inferred as surrounding and influencing it). The respective distinction can generate a beautiful interplay. Brassaï has produced a few such instances; in one of his works (entitled *Toilette dans une maison de passe, rue Quincampoix*, around 1932 – Fig. 4), the objects within the image are arranged according to different depths of field: in the background, there is a wardrobe (forming the hypotenuse connecting the legs that are the room's walls), with mirrors attached on its two doors; one of the mirrors shows the reflection of a naked woman, with her back turned to it (she is getting dressed); the other door is hidden by the man in the foreground, turning his back to the photographer, and setting his clothes straight – from that angle, the photo cannot show his face; we can postulate two more distinct planes outside the visual field: the position of the photographer, behind the man, but at such an angle that allows him to see the reflection of the woman, and another plane outside the visual field, within the visual universe (behind the photographer), where the woman is.

Fig. 4

We may not know French (which would allow us to understand the caption), we may not know that some hotels received prostitutes and their clients. It is not even important to know it. What matters is that the photographer does not have his characters occupy the same space (even in different positions or orientations), but leaves them in different spaces, turning their backs to himself and the beholder. It is precisely this arrangement that prompts us to consider the plane where the photographer stands, among the many planes within this photograph.

Immediately after lovemaking, the partners no longer seek eye contact; there follows a brief isolation, that is, a return to their previous state. But the persons are still together in this space space. The mirrors on the wardrobe door in Brassaï's picture prompt the viewer to compartmentalize this space, destroy its unity, estrange the characters even more. Indeed, the beholder must contemplate the image for a few moments in order to reconstruct the spatial unity of this room.

According to David Hockney, however, photography is not in the best position to handle space: The camera sees everything at once – in one click of the lens, from one vantage point. Painters, according to Hockney, don't. And the fact that painters need time to see everything is what creates space. One might say that the reason photographs cannot represent space is that there is no space at the moment when the camera captures the picture (Hockney, Gayford, 2017,pp. 83–84).

Elsewhere he states that the photograph – and the camera – cannot handle space, but only surfaces (Hockney, Gayford, 2017, p. 206). I understand from Hockney that in order to create space, the painter needs the time to reflect on it, on what it looks like, and then on what it will look like on the canvas; thus there is a space in, and of, the surrounding world, and a space of the visual artifact, which the painter can make more or less evocative of the space exterior to the respective artifact. By contemplating this space for a while (through sight or in the mind) the painter establishes selections and hierarchies and then transposes them on the canvas (it should be noted that Hockney fails to explain in detail how these selections and hierarchizations are evinced by the painting). According to Hockney, photographers are unable to do so; they "see everything at once"; so the filtering process is missing. But, one may wonder, can't the photographer map out the space by tapping into the possibilities of displaying it, as they are offered by the relative posi-

tions *within the respective space*? How does the photographer select one of them?[3] Can't one construct a space by defining those shapes that serve as landmarks in the unfolding of that space?[4]

In my opinion this is an issue that has to do with the concern (a constant one – for some commentators) with defining a radical separation between the cultural and the biological dimensions of vision. Already the distinction between figure (form) and background (see their correlation in Groupe μ, 1992, p. 68) results in constructing a space:

> It is a *figure* what we submit to such attention that engages an elaborate cerebral mechanism for local scrutiny (…) It is *background* what we do not submit to this type of attention, and for this reason is analyzed by less powerful mechanisms for the global discrimination of the *textures*. The effects of this opposition are well-known. Let us mention two, not without importance in establishing the visual codes:
> 1. the background pertains to the nature of the field, is indiscriminate and by definition limitless;
> 2. the background appears to exist *beneath* the figure which, from then on, appears to be closer to the viewer than the background is (Groupe μ, 1992,pp.67-68).

In Brassaï's photograph, the successive margins framing the characters (the frame of the photo, the sides of the wardrobe, the mirror edges) and the different sizes of the two living figures (the man and the woman) create visual premises that are equally biological and cultural, to differentiate and isolate the figures.

Equivalents, the series of photographs produced by Stieglitz in the 1920s and early 1930s, provides us with useful conclusions on the nature of framing and what is at stake in framing. These are photos of the cloudy sky. As long as a structure (a composition) cannot be identified–that is, as long as the photos postulate the absence of a basic composition– the beholder is seized by a sense of extreme confusion (Krauss). Stieglitz' pictures of clouds demonstrate that photography no longer is what we always believed it to be–namely, the possible extension of the experience of our material presence in the world. To Rosalind Krauss, in these photographs, framing is by no means a mere mechanical phenomenon. It is the only thing that constitutes the image, and in constituting it, implies the fact that photography is an absolute transformation of reality (Krauss, 2013, p.163).

Thus, framing opens a new direction in the discussions on photography: Kracauer has pointed out that photographic expressivity is predicated upon a "precarious balance" struck between the inevitable penchant for realism and the formative tendencies that are born out of the photographer's internal, psychological dynamics (Trachtenberg, 1980, p.260). This brings us to the issue of composition, which can be placed among the formative endeavors that the photographer undertakes and the photograph reveals.

> We must rid ourselves of a confusion brought about by continually comparing photography with the fine arts. Every handbook on photography talks about composition. The good photograph is the well-composed one. Yet this is true only in so far as we think of photographic images imitating painted ones. Painting is an art of arrangement: therefore it is reasonable to demand that there is a kind of order in what is arranged. Every relation between forms in a painting is to some degree adaptable to the painter's purpose. This is not the case with photography. (Unless we include those absurd studio works in which the photographer arranges every detail of his subject before he takes the picture.) Composition in the profound, formative sense of the word cannot enter into photography (…) The true content of a photograph is invisible, for it derives from a play, not with form, but with time (…) a photograph bears witness to a human choice being exercised. This choice is (…) between photographing at X moment or at Y moment (Berger, 2013, p.19).

Again, things are more complicated than they seem. There is no single type of composition (usually a line of argument begins with a truism); as art history teaches us, disciples would learn the manner of painting from their masters; then the time came when some of them took the old techniques to a level of expressivity deemed the ultimate perfection, or when others produced new ones. On a visit to San Luigi dei Francesi made by Federico Zuccari (the

[3] This point provides the reason I have chosen to discuss the photographs by Michel and Brassaï.

[4] Note that, in photography, the setting of diaphragms will define, with various degrees of accuracy, the respective shapes and will render transitions between space portions more or less abrupt

first president of Accademia di San Luca) in order to see Caravaggio's works, in response to the interest stirred by the respective works, Zuccari remarked: "What is all this commotion about? I can see nothing but Giorgione's manner of thinking." , Giovanni Baglione, a stark opponent of Caravaggio and later his biographer,

> could not conceal his fear that Caravaggio's works were actually so shocking, so unexpected that powerful painters such as Zuccari attempted to disguise their concern. The master's mocking words betray the troubled reaction of an entire generation, which erected the wall of mistrust in order to defend itself against a novelty it could not comprehend (D'Orazio, 2013, p. 21).

There are various manners of bringing order to figurative paintings, and transition from one to another is not exactly easy. Moreover, these manners become exhausted and have to be replaced (the exhaustion of operations inevitably and dramatically results in works deprived of any power of expression):

> Their paintings (of history painters such as Paul Delaroche, Horace Vernet, Léon Cogniet, Robert Fleury; history painting was the highest genre in 19th-century France – C.P.) stand out due to the lack of any personal quality. When you look at one of their paintings, whatever its subject, you are first struck by its immediate intelligibility. The work lends itself to the viewer entirely, down to the smallest details. In the compositions of Léon Cogniet, Delaroche or Vernet, the drawing complies with pre-established proportions. The color is treated similarly. Whether the brushstrokes are light or heavy, the painter's dominant intention is to be veridical. These principles exclude all intermediate hues. When he *intends to be serious*, the painter fills the painting with all sorts of details. If he intends to be *cheerful* or *dazzling*, he multiplies the colors (Freund, 1974, p.63).[5]

What order did Berger expect to find in photographs? Would he have been satisfied with any of the types of pictorial order? Can we expect photography to have produced (and to continue to produce) formulas, rhetorical devices, or schemes whose variations are capable of generating meanings? You cannot achieve a pattern of order when you lack the freedom to handle your own means of expression, the forms you devise by comparing what you have learned to do with what you aim to achieve. There is order in photography, but we should not seek it among the codes of order produced in the art of painting throughout the times (or rather, we should finally go beyond them). Seeking these patterns of order, in my opinion, should mean seeking modes which, for lack of a better word, I would term 'prepictorial' modes of organizing visual fields.

There have been attempts at describing such forms of order. In *The Power of the Center*, Rudolf Arnheim aims to present

> a compositional scheme common to works of visual art of whatever time or place – a condition that needs to be met by all art if it is to fulfill its function. How would this universality come about? Obviously, there is no convention prescribing one particular form of composition to all cultures. Rather, its principle would have to be deeply rooted in human nature and ultimately in the very makeup of the nervous system we all have in common. (Arnheim, 2009, pp. 1–2).

This scheme must be located at a greater depth of perceptual organization than the culturally devised schemes that shape perception (schemes discussed, among many others, by Wölfflin, for instance). Arnheim resorts to two pairs of principles in the interpretation of composition: centricity and eccentricity, geometric and energetic. The principles are derived from spatial relations, manifest in visual forms. But the richness of the physio-neuro-psychological interconditionings of the schemes is much greater, and their articulations are insufficiently known yet. As long as it does not own its expressive material, photography has to look for its specific character in orderings which as of yet are un-systematizable, because they are unpredictable: the material to be ordered a) changes all the time, b) is greatly resistant to ordering. The eye incessantly re-orders and reconstructs orders, but it seldom succeeds in achieving them. Not surprisingly, Edward Weston insisted polemically in favor of another, different approach to art than "the rules of composition":

[5]Gayford tells us that Poussin was known to employ a sort of miniature theatre in order to illuminate and compose his painting: in other words, to stage the narrative (after Hockney, Gayford, 2017, p.167). Similar situations are mentioned in the same source.

> When subject matter is forced to fit into preconceived patterns, there can be no freshness of vision. Following rules of composition can only lead to a tedious repetition of pictorial clichés. Good composition is only the strongest way of seeing the subject (Weston, in Trachtenberg, 1980, p.175).

I understand Weston as cautioning against preconceived compositional schemes to ensure photography avoids subordination to painting. However, it is a long way from merely avoiding subordination to successfully revealing the importance of a subject. Photography, though, is not the only form of visual art that can fail to achieve this (even though the terms of this exposure to risk differ from one category of visual artifacts to another). Krauss wrote that resorting to the ontological argument may affect the way in which a medium (photography, painting, theatre, etc.) is considered as an art. Defining a category a priori – here, defining a photograph in relation to the way it operates in framing reality – may give the impression that this category has always existed (and was waiting to be noticed, brought to fruition, and explored). In defining a category a priori, we obscure a much more important aspect of its nature, that of the risk inherent to the creation of any work of art–or at least any work considered o 'high art'. Often, this risk lies in working intuitively, without the slightest certainty or guarantee of success (Krauss, 2013, p.164).

We may achieve a better understanding of photography by reflecting more thoroughly on the risks (mentioned by Krauss) involved in photography. Aleksandr Rodchenko is one of the first major photographers to experiment with new ways of looking at things; the following statement is reiterated in several of his discussions on photography: "One has to take several different shots of a subject, from different points of view and in different situations, as if one examined it in the round rather than looked through the same key-hole again and again." What could prevent us, in this case, from acknowledging the photographer's privilege of recognizing around him what is meaningful enough to become part of a balanced composition? Painters seek the path toward this balance and sometimes reach it. The photographer also undertakes such training and, in his turn, may discover in his proximity this or that configuration that looks ready and waiting to be discovered (he has noticed it, and presents it to us); or he may create a configuration himself, with the materials available (see below, the photo taken in India by Cartier Bresson).

According to David Hockney, a painting lets us perceive the time invested in producing it. This time is not uniformly distributed across the surface of the painting. In contrast, in photographs, the same amount of time is allocated for every point of the surface. This is why one cannot look at a photograph for too long, says Hockney. It is actually taken only in a split second, so that one cannot have a layered perception of the subject. The portrait Lucian Freud made him required one hundred and twenty hours of sitting, and the beholder can see all that time *laid* layer after layer within the painting. This is why, Hockney believes, it is infinitely more interesting than a photograph (Hockney, Gayford, 2017, p.78). Gayford has a similar explanation for the value of painting: he had met Madame Ginoux five months before she sat for him, and, during this time, must have watched her for hours–a substantial expenditure of time that was integrated into the painting (Hockney, Gayford, 2017, 78).

It is possibly true that producing a painting requires a greater investment of time than producing a photograph. But since when is the time spent producing a visual artifact an essential axiological criterion? In contrast to Gayford and Hockney, Van Gogh boasted about having completed *l'Arlésienne* in only one hour. I have no idea whether Smith knew doctor Ceriani well enough to be able to portray him adequately, but *Dr.Ceriani with Coffee Cup* (1948) certainly conveys more than the portrait made by Lucian Freud: the physician has just completed a surgical intervention; he is resting one elbow on the kitchen counter, with a cup of coffee in one hand and a cigarette in the other. He is clearly exhausted and is staring vacantly. His medical scrubs, as well as the expression on his face and his body language, stand in sharp contrast with the kitchen's atmosphere and furnishings. Freud's portrait shows a concerned, upset Hockney, also staring absently (above his glasses). But his inner life, which we may reconstruct by the exterior appearance of his body and face – one of the criteria for the assessment of a portrait (Arnheim, 1992, pp.105-106) – is more impressive in the photograph than in the painting, the amount of time spent on creating each notwithstanding.

The photographic composition, resulting from the concern with "the strongest way of seeing the subject", is the effect of a permanently alert state:

> Photography is a new type of art, based on the instant; we work in motion, with something of an intuition of life, and in motion the photograph has to strike the expressive balance. Our eye has to measure, to evaluate continuously. We produce overlappings of the lines simply by moving our head by half a millimeter, but this can only be done at the speed of a reflex movement, and luckily spares us the attempt of doing 'Art' (Cartier Bresson, p. 3).

A photo taken by the French artist in India (Madura, Tamil Nadu), in 1950, shows a woman cradling her baby (Fig. 5). A cartwheel can be seen very close to her, and takes up about one quarter of the whole picture, touching its right and upper margins. The woman is covering the child's head with her left palm; her fingers are spread apart, so as to better protect him. The wheel spokes and the woman's fingers are juxtaposed with the ribs of the malnourished child. What we can see is the threat of doom. Probably the distance between the two circular shapes (the wheel and the baby's head) was larger than it appears in the photo. With the material provided by this portion of a street (and seemingly to make his point) Cartier-Bresson managed to construct a rich and powerful configuration with its own internal order.

Fig. 5

Fig. 6 and Fig. 7 further demonstrate how the same material found in the world can be employed to construct different evocative, artistic configurations in photography.

Fig. 6

Fig. 7

The limits set for the visual field, and the configuration arranged within it, depend on the visual material encountered. Very often people remark that many photos include elements that they would be better off without, since their presence disrupts the coherence of the artifacts' connotations (we deplore the fact that the author could not avoid them). Structural semantics has termed *isotopy* the homogeneity of a given layer of meaning, or level of significance; isotopy is a prerequisite in understanding a text. In principle, 'classic' works (writings, paintings, sculptures, etc.) have high degrees of isotopy. A newspaper, a pipe, and a glass of whisky form a configuration that is understood much more easily and quickly than that of an old man wearing a folk costume, an electronic cigarettes vending machine, and a tsunami. However, photographs often include elements which upset the isotopy, and without which the picture would be perceived as more gratifying.

> However selective photographs are, they cannot deny the tendency toward the unorganized

> and diffuse which marks them as records. It is therefore inevitable that they should be surrounded with a fringe of indistinct multiple meanings. (To be sure, the traditional work of art carries many meanings also. But due to its rise from interpretable human intentions and circumstances, the meanings inherent in it can virtually be ascertained, whereas those of the photograph are necessarily indeterminate because the latter is bound to convey unshaped nature itself, nature in its inscrutability. As compared with a photograph, any painting has a relatively definite significance. Accordingly, it makes sense to speak of multiple meanings, vague meaningfulness, and the like only in connection with camera work) (Kracauer, Trachtenberg (ed.), 1980, p.265).

The terms *unshaped nature* and *inscrutability* indicate the (rather frequent) difficulty faced by photographs in creating meanings coherently enough for the mind to recognize them easily and formulate them clearly. Kracauer refers to *the traditional work of art*; but what about those works of art which are not traditional, and whose meanings are hardly lexicalized? What about those who look for beauty in "the chance encounter of a sewing machine and an umbrella, on an operating table" (Lautréamont)? This aesthetic domain is seldom evoked when critics attempt to explain why photography can never be like painting;and the fact that photography can achieve its own substantial expressivity is rarely addressed at all. In other words, critics maintain that, by showing the world as it really is, photography falls short of isotopy. This demonstrates a tendency to ignore the avowed intention of some photographers (especially cityscape photographers) to reveal and emphasize (the seriousness of) the lack of coherent significance that characterizes the world. Clearly, disorder is shocking. There are too few who appreciate this (apparent) lack of coherence, and who find ways of harmonizing and taming the tangled meanings of the modern world. Among them are the surrealists. They were attracted by the unshaped (that is, by the unrecognizable shapes);. To the unshaped, Bataille assigns the task of deconstructing formal categories, of denying that everything has its "own" form, and of imagining meanings deprived of a definite shape, like a spider or worm crushed under the foot (Krauss, 2013, p. 205).

And Brassaï's *Nu, la poitrine*, a photo dating from 1931-32, displays the unshaped by having the protagonist tilt her head backwards, and thus reducing it to the chin, transforming the breasts into horns, and the torso and illuminated arm into a face and an ear.

Between the frustration generated by those photographs that fail to reach a high degree of isotopy, and the surrealist intention of destroying the identity of things by altering their formal aspects (those aspects which best enable us to categorically recognize things) there seems to be no connection. However, I believe that both are ways of discovering the dynamic plasticity of the world–its protean nature, which is constantly capable of causing confusion. In both cases, we are unable either to recognize the objects, or to perceive them as part of a general network of any satisfactory significance (to this point, Krauss noted that the surrealists used flea markets as miniature representations of the city and of the extraordinary diversity of objects constituting the material grounds based on which an infinite number of references can be predicated (Krauss, 2013, p.184)).

After all, maybe we ought to see the isotopic deficit of photography in a positive light. It helps us first to understand that meaning is hard to achieve; in photography, to become able to clarify and emphasize meaning presupposes a process of selecting and ordering the elements of the surrounding world. The elements perturbing isotopy may reveal how the photographer has perceived the world and derived meaning from it, a process that is cannot always be ordered or completed. Second, it helps us to understand that not all the meanings we produce are highly formalized–for example, the artifacts belonging to the high arts (see the pictorial genres: portrait, landscape, history painting, etc.…),nevertheless, they still are expressive landmarks in our existence.

However, there is more to photography than its insufficiently orderly character, which cannot completely free itself from the objectness of the world, and therefore, from its possibly deficient intelligibility. Kracauer himself points to the need for photography to seek a "precarious balance":

> What counts is the 'right' mixture of his (photographer's) realist loyalties and formative endeavours – a mixture, that is, in which the latter, however strongly developed, surrender their independence to the former (in Trachtenberg (ed.), 1980, p.260).

The formative tendency, then, does not have to conflict with the realistic tendency. Quite the contrary, it may help substantiate and fulfill it (…) the

photographer sees things in his 'own soul' (Trachtenberg (ed.), 1980, p.260).

Here, I perceive a dialectical relationship: in the photographic artifacts we identify elements extracted from reality with astonishing precision, but in those instances that should be crucial in defining the expressive potential of photography, we can't help finding in the manner of configuring these elements, evidence of the active, intentional presence of the authors of the respective artifacts. The fragment of reality and the formative intention of the photographer who captures it are inseparable:

> The eye outlines the subject, and the camera does its work, fixating on film what the eye has selected (…) To me a photograph is the simultaneous recognition, achieved in a split second, of a fact's significance on the one hand and on the other, of the rigorous organization of visually perceived forms which express that fact. (Cartier Bresson, p.3).

Fig. 8

In 1986, Jean-Marie Floch analyzed a famous photograph taken by Robert Doisneau, *Fox-terrier sur le Pont des Arts* (Floch, 2018, pp. 31–41) (Fig. 8). The picture also circulates in a version where on the viewer's left and in the background, a dark figure is about to leave the visual field (Fig. 9). Noting that such trimming is common, Sorin Alexandrescu (2008, p. 9) has tried to account for the presence of this person in the picture. In his opinion, on the one hand this figure emphasizes the diagonal line of sight along the bridge, and on the other hand he and the dog are the only ones not interested in the scene in the middle ground (an artist is painting a nude, while the canvas is only partly visible; behind the painter who is hiding her from view, a woman – possibly the model – is sitting on a bench: we identify one shoe-clad foot; in the foreground, another passer-by is watching the scene, slightly bending in an attempt to look at the woman; his dog is staring at the photographer/ beholder).

Fig. 9

According to Floch, the dog who is on the left is blocking that side of the bridge from where the photographer (and the beholder, as well) could have caught a glimpse of the woman (ata different angle). Alexandrescu accepts this explanation and points out that looking from the spot occupied by the dog, one might have assessed how the elements of reality (the model) match their artistic transposition (the nude). He also notes that the only ones to direct their gaze outside the visual field are the dog and the remote passer-by on the left – who has reached the far end of the bridge by walking past the photographer – and whom he considers uninterested in the scene. Finally, the passer-by on the left is part of the context – close to the scene, but indifferent to it (this distinguishes the man from the dog, who according to the two commentators is not detached from it, but stands guard and prevents visual access to the scene[6] – once

[6] Actually the dog is not blocking access to the spot from where the model can be seen, but is reacting to the photographer coming too close. However, the original explanation of its behavior, not related to the photograph, is altered by Doisneau who discusses its position and stance in relation to those of the other characters. By choosing the framing and the vantage point, any photographer can alter the significance of the scene inspiring him.

again, Alexandrescu had noted that these two characters are looking outside the visual field). The dog appears in order to "express an interdiction, a ban to be observed by the artist: *in art, look not for what is true, but for what is meaningful"* (Alexandrescu, 2008, 29). The Romanian semiotician sees this photograph as a meditation on veridiction: the dog that does not allow the photographer to see the woman on the bench (is she naked? is she not?[7]) offers an alibi to the photographer who intends to demonstrate that a photograph does not necessarily guarantee the truth (28). Earlier (19), Alexandrescu noted that Doisneau's photo includes almost all types of beholders, and thus it is an allegory of vision itself.

In his postface to Floch's book, Dario Mangone informs us that Doisneau and the painter, named Daniel Pipard, were friends and that the photograph had been designed by them (Floch, 2018, p.118). Maybe the dog's owner was another friend of Doisneau and Pipard and was persuaded to take part in producing this artifact, together with his dog.[8] But how about the person about to leave the visual field, on the left?

As one among almost all types of beholders (according to Alexandrescu), the presence of the remote passer-by on the far left is justified in the allegory-picture. Why, then, is he excluded from some of its reproductions? Because, on the one hand, the allegorical character is not recognized by everybody and on the other hand, because his position (the silhouette touches one of the margins) and size (very small) seem to deny the pedestrian the ability to participate in the significance – whatever it is felt to be – of this photograph.

This leads us to the idea that between representing fragments of reality and employing them with a view to the success of the photographer's "formative endeavors", the balance often becomes "precarious", as Kracauer puts it (Doisneau did stage quite a lot of his photos, some of them so ambiguous that they stirred scandals).

Of course, the formative pursuit of photographs is not always highly ambitious; there are not many pictures that raise the issue of veridiction. The *Fox-terrier sur le Pont des Arts* deserves our attention for two reasons: it has revealed the author's highly ambitious formative aspirations, and it has demonstrated the commentators' interest in a higher degree of coherence of images: pictures are expected to be perfectly integrated compositions (even those items that seem to be alien to the isotopy of the image must find their place within it, because flawed isotopy is disturbing).

We may draw a brief conclusion: there are several explanations for the semantic imprecision of photographs. For instance: they contain shapes incompletely integrated isotopically, which the photographer was unable to avoid; or the photo was simply not taken at the right moment–the moment carrying the strongest force of expression, and rich enough to allow beholders to imagine the previous and following moments.

Fig. 10

However, there are situations where this imprecision becomes fascinating. In a 1945 photo taken by Paul Strand (*Susan Thompson, Cape Split, Maine*) (Fig. 10), a woman is standing on the threshold of a room that seems to be a kitchen. Over her dress, she is wearing a white crumpled apron (how many times she has wiped her hand dry with her apron, for lack of time to find a towel?). Down, on her right, is a tall lidded pail. Susan is about to move her arms; her eyes are not yet set on any object that might stir her. A kitchen – with so many chores to do; a pail, a wrinkled apron – proof of her hard work; a door

[7] Other photographs reveal that the woman was in fact wearing a trench coat. But the rhetorics of this image requires us to ignore this.

[8] "In the Paris of the 1930s, Brassaï like many others working in the same field, was known to 'stage' his most important pictures, by paying models or asking his friends to act in scenarios he had either learned about, or invented himself" (Bate, 2016, p.73, quoting from Grierson).

frame – separating two distinct spaces; a pair of arms – perhaps announcing,, a movement that we cannot anticipate; a brooch – telling us that Susan (a true woman after all!), can be self-aware, too; the gaze – inquiring into another world.

It is Cartier Bressson who theorized about the *decisive moment*, but many photographers applied his views *avant la lettre*. Paul Strand described the decisive moment in reference to this image as Susan suspending her work and communing with a different universe than that of everyday chores. The door frame, or the frontier marked by the wooden post, separates not the inside of the kitchen from the outside of the Cape Split yard; rather, it detaches Susan's world of everyday toil from another world–a world that, even relegated to her imagination, is a miracle. Susan Thompson never really left the former to join the latter, but this image of her conveys a unique relationship to both the space she occupies and the space she dreams of. This runs contrary to the claims of David Hockney, who states that photographs are unable to render space (Hockney, Gayford, 2017,p.83), and Susan Sontag, who points out that "it is in the nature of a photograph that it can never entirely transcend its subject, as a painting can" (Sontag, 1977, p.95).

In the literature on photography, there is frequent commentary on the idea that despite their interest in being selective, photographic artifacts are disorganized and disorderly, carrying many unclear connotations (see *supra*: Kracauer, in Trachtenberg, 1980, p.265). This tenet certainly deserves greater and more nuanced attention. Photography is judged according to criteria derived from the high art (criteria which do not concern the world of art in its entirety, even though that portion of the arts they apply to is most often defined and acknowledged by society as 'great' art). This approach tends to obscure the originality of photography's means of expression (showing *how* it can process the protean material of reality–often deprived of clear significance–to produce configurations whose expressive-symbolic load is able to shape a wide range of experiences – moral, aesthetic, etc., for its beholders).

*

Thus, one can list the characteristic features of photography which pertain to its power of expression. The accuracy of definition, and the broad spectrum covered from the most gradual variations in hues, brightness, and saturation, to their radical contrast, were remarked from its very outset. Framing is an essential expressive tool; it often creates an interplay between the visual field and the visual universe (what is beyond and outside this field, and can have an impact on its significance). Its persuasive force reinforces and is reinforced by the vantage point, the spot from which the author of the visual artifact regards the scene; this can generate a wide array of angles and depths. Finally, photography seeks to strike a balance, which Kracauer describes as precarious, between realist loyalties and formative endeavors, which can gain the recognition of its power to transpose and symbolize.

I have endeavored to provide a nuanced overview of the issues pertaining to the status of photography with regard to its expressive possibilities. Of course, the issue of its contents also needs in-depth examination (conceptual photography, documentary photography, etc.). To Barthes and Sontag, photography is elegiac, a *memento mori*, etc, and to others photography is a preferred medium for interfacing with different concepts, ideas, and feelings.. The debates on the artistic quality of photography have in their turn outgrown the exchange of more or less axiomatic assertions. As photography continues to grow in prominence in our society, it definitely deserves more thorough theoretical investigation.

References

Alexandrescu S (2008) Le mystère du fox-terrier, in *E/C*, www.ec-aiss.it. (https://ecitydoc.com/download/le-mystere-du-foxterrier1_pdf)

Arnheim R (2009/1988) *The power of the center. A study of composition in the visual arts.* Berkeley: University of California Press.

Arnheim R (1992) Caricature. The rationale of deformation. In *To the rescue of art*, 101–114. Berkeley: University of California Press.

Barrett T (1996/1990) *Criticizing photographs. An introduction to understanding images.* Mountain View CA, Mayfield.

Barthes R (2005/1980) *Camera luminoasă. Însemnări despre fotografie*, rom.transl.. Cluj, Idea Design & Print

Bate D (2016) *Photography. The key concepts*, London: Bloomsbury.

Berger J (2013/1968) Understanding a photograph. In *Understanding a photograph*, 17–21. New York: Penguin.

Berger J, Mohr J (2016/1982) *Another way of telling. A possible history of photography.* London: Bloomsbury

Bourdieu P (1996/1965) *Photography. A middle-brow art*, engl.transl. Cambridge: Polity.

Cartier Bresson H (1952) L'instant décisif, foreword to *Images à la sauvette* (https://www.zintv.org/IMG/pdf/Henri_Cartier_Bresson_-_l_instant_decisif_.pdf)

Clarke G (1997) *The photograph.* New York: Oxford University Press.

D'Orazio C (2013) *Caravaggio segreto*, Segrate MI: Sperling & Kupfer.

Floch J-M (1985) *Petites mythologies de l'œil et de l'esprit. Pour une sémiotique plastique.* Paris Amsterdam: Hadès Benjamins.

Floch J-M (2018/1986) *Forme dell'impronta*, it.transl. Milano: Meltemi.

Freund G (1974) *Photographie et société.* Paris: Seuil.

Groupe μ (1992) *Traité du signe visuel.* Paris: Seuil.

Hockney D, Gayford M (2017/2016) *O istorie a imaginilor*, rom.transl. Bucureşti: Pandora M.

Kracauer S (1980) Photography. In.), *Classic essays on photography,* ed. Alan Trachtenberg, 245–268. Stony Creek CT: Leete's Island Books.

Krauss, R (2013/1974) L'impressionnisme: le narcissisme de la lumière. In *Le photographique. Pour une théorie des écarts*, fr.transl., 77–90. Paris: Macula.

Krauss R (2013) Stieglitz: Equivalents (1979). In *Le photographique. Pour une théorie des écarts*, fr.transl., 153–164. Paris: Macula.

Krauss R (2013) Les noctambules (1980). In *Le photographique. Pour une théorie des écarts*, fr.transl., 165–184. Paris: Macula.

Krauss R (2013/1984) Corpus delicti. In *Le photographique. Pour une théorie des écarts*, fr.transl., 197–235. Paris: Macula.

Kress G, Van Leeuwen T (1996) *Reading images.* London: Routledge.

Maynard P (1997) *The engine of visualization. Thinking through photography.* Ithaca NY: Cornell University Press.

Newhall N (ed.) (1975) *Edward Weston, the flame of recognition: his photographs, accompanied by excerpts from the daybooks and letters*. New York: Aperture.

Sontag S (1977) *On photography.* New York: Picador.

Weston E (1980) Seeing photographically. In *Classic essays on photography*, ed. Alan Trachtenberg, 169–175. Stony Creek CT: Leete's Island Books.

BALKAN JOURNAL OF PHILOSOPHY. Vol. 11, Issue 2, 2019

INTERPRETING THE CONCEPT OF POLITICS IN TERMS OF PREFIXATION

Irina Zhurbina

Udmurt State University
Universitetskaya str., 1/4, Russia, Izhevsk
e-mail: soloveyIV1@mail.ru

Abstract:
The article analyzes the revision of the concept of politics caused by the exhaustion of the ideological paradigm. In modern philosophy politics acquires new meanings through prefixing, resulting in the emergence of such concepts as archi-politics, para-politics, ultra-politics, trans-politics, or bio-politics. These new concepts close the philosophical source of politics laid by the Greek tradition. The departure from philosophy as the source of politics is completed with the idea of police, in which prefixing as a way of conceptualizing politics reaches the linguistic limit. However, modern philosophy encompasses a more positive attitude, which is linked to the hermeneutic tradition of philosophizing of Heidegger and Gadamer that focuses on the preservation of thought and language in the source of political existence.

Key words: Archipolitics, Parapolitics, Ultra-politics, Transpolitics, Biopolitics, Metapolitics.

In the post-modern era, the existence of the world as a whole is most fully expressed in the concept of globalization. Global processes are changing the essence of the social world, which is losing its structure, stability, and guaranteed order. We come across the effect of 'escaping' social reality, when it becomes impossible to state anything about the social world using the terminology of traditional discourses. The era of global change allows us to speak about 'the end of the world as we know it' (Wallerstein, 1999). There is a widespread belief that in an era of global transformations, no idea that claims to be the dominant ideology can any longer perform the function of a political strategy governing society. All ideologies that were more or less important – communism, liberalism, conservatism – have reached their limits and have become unsustainable. The world is entering an 'after liberalism' era, which will inevitably change the general paradigm of politics (Wallerstein, 1995).

In philosophical discourse, to overcome the ideological paradigm in politics one has to re-think the source of politics set by Greek philosophy. The sources inevitably refer to the Greek concept of ἀρχή, which has two meanings. It means both 'beginning' and 'command, rule', which indicates that the source is always the dominant principle (Agamben, 2013). Since the source of politics is Greek philosophy, referring to the source involves further interpretation of the word πολιτικά. Noteworthy is the fact that new meanings behind the concept of politics evolve out of prefixing as a linguistic way of conceptualizing. Currently, the most important concepts are biopolitics (Foucault, 2008), transpolitics (Baudrillard, 1983), metapolitics (Badiou, 2005), parapolitics, archipolitics (Rancière, 1999), and post-politics and ultra-politics (Žižek, 2009), which do not form either a conceptual unity or a holistic view. Instead, they show different modes of politics' existence. The increasing number of prefixes indicates that in the socio-philosophical discourse, the concept of politics begins to live its own life as a linguistic construct. It should be noted that the concept of politics is fleshed out with content only when it has a prefix. In prefixal interpretation, politics becomes a homonymous notion with a formal structure. The word politics, pronounced and written in the same way, is each time re-read by researchers through a lens established by the prefix with which the word starts.

Prefixing is done by adding to the root word of 'politics' a Greek or Latin prefix bio-, archi-, ultra-,

trans-, para-, meta-, or post- through which the word takes on an additional meaning, which is not so much the emergence of a new one as the revision of the old one. There is a steady trend of using only those prefixes which expand the meaning of the word 'politics'. Prefixing constrains the formation of new concepts by transforming the old ones. The general meaning of prefixes is determined by semantics, according to which the use of such prefixes as super-, post-, neo-, hyper-, meta-, or archi- makes it possible to go beyond the traditional meaning of the word. Prefixing violates the tradition of using the concept of politics that exists in the language of philosophy. The meaning of politics is expanded so as to acquire a variety of meanings, i.e. lose the terminological stability of a category. One can clearly see how prefixes that are synonymous in terms of their form (archi-, ultra-, trans-, meta-, post-, para-) with a similar meaning of 'going beyond' begin to differ conceptually. Having explored the prefixes, we can distinguish their temporal meanings through which it is possible to see the concept of politics in perspective – the beginning (archi-), the middle or process (bio-), and the end (post-). While archi- returns us to the revision of the past tradition existing in permanence, the prefix post- seems to turn us to the future of politics, which can be defined and redefined endlessly because of its uncertainty. An attempt to revise the past tradition gives rise to various ideas of the beginning of politics, and an appeal to the future returns it to the present in the 'end of politics' concepts. Between the beginning and the end, as the past and the future, unfolding is the space of the politics of life determined by the concept of biopolitics.

The beginning of politics opens with Plato's theory of state, politics, and laws. The modern concept of archipolitics, according to Rancière, returns us to Plato's "political idyll of achieving the common good by an enlightened government of elites buoyed by the confidence of the masses" (Rancière, 1999, p. 93). The prefix archi-, meaning, firstly, precedence, rule, and, secondly, a high degree of anything, indicates that politics is subject to a higher principle. In the teachings of Plato, the source of politics, its ἀρχή, is knowledge understood as λογος. Knowledge is the highest Good, i.e. the principle without which natural life (physis) exists in its non-completeness and randomness as a chaotic multiplicity of individuals deprived of their unity. Philosophy, which is the highest degree of manifestation of knowledge in general, transforms politics into art, which fully presents itself in the implementation of the law (nomos) of Good, Kindness, and Justice in state government. State organization depends on how good the politicians' knowledge of the highest ideas of Good, Kindness, and Justice is.

The concept of archipolitics departs from the philosophical source of politics. Knowing loses the meaning of logos (λογος), which leads to the ultimate rationalization of not only thinking, but also the entire sphere of knowledge. Rationalization becomes a way of organizing the social world, in which science and technology gradually begin to dominate. Knowledge in politics is necessary to the extent to which it contributes to rationalization and optimization of the governance of the social world. With regard to modern conditions, archipolitics is a concept justifying the power of science and expert management of the state. This situation occurs when the general understanding of political power changes. It is as if archipolitics demonstrates the transfer of power from politicians to 'experts'–bearers of scientific knowledge. In the context of this general transformation of ideas about political power a special form of government – epistemocracy – appears. Epistemocracy is a result of the development of management techniques, demonstrating the level of human knowledge. Rationality becomes the basic principle of state governance. It establishes a connection between politics as practice and politics as knowledge, with science being its supreme incarnation. Epistemocracy changes the relationship between science and politics. Within epistemocracy as a special form of government, an intellectual generalist independent of the state is replaced by an intellectual specialist, who is politicized and partisan, and involved in the state management process (Foucault, 1980, p. 129). Specialists get involved in the vicissitudes of political life, where they play the role of experts, allowing them not only to legitimize their own scientific knowledge, but also to use the state to fulfil their own power ambitions.

The power of scientific knowledge divides society into two non-uniform parts: those knowing and not-knowing. Citizens take on an obligation to be controlled, i.e. agree to be not-knowing and act as objects of control. The not-knowing part of society, due to their incompetence, i.e. lack of expertise, are, in fact, removed from discussion and decision-making that affect the development and well-being of society. The knowing part is represented by experts. The role of the expert is determined by the presence of specialized knowledge sufficient for

decision-making in the field of economics, law, education, medicine, science, information technology, etc. The level of professional qualifications or expert competence determines the significance of the issues discussed with the expert. Expert systems are developed by means of specialising in certain areas of management, resulting in narrowly specialized expert groups dealing with specific issues of ecology, chemical and biotechnological production, nuclear energy, genetics, etc. There is a peculiar pattern: the narrower the specialization becomes, the more limited the scope of the expert knowledge is and the bigger the number of experts becomes.

Expertocracy can be regarded as one of the ways epistemocracy manifests itself. It is believed that an expert opinion should increase the predictability and controllability of natural, technological, and social processes. However, no expert opinion can either give a definite answer to the question about the long-term effects of any event that has occurred as the result of a natural or man-made disaster, or predict the ways in which human behavior may engender or contribute to a potential accident. Therefore, any expert solution increases the degree of uncertainty and creates unforeseen and unintended consequences. There is always a chance that something which experts believed to be impossible will happen. Expertise includes risk assessment. Unintended consequences are no longer an exception, but an inevitable rule of the system of government. Rationalization of the expert government system reaches its peak in situations where it changes its priorities and strategies–where, for instance, the government transitions from ensuring peace, security, and stability to focusing on reducing the likelihood of damage and environmental, chemical, nuclear, genetic, or other threats. Risk management is intended to include risk experts for whom risk assessment itself is a very risky activity. Risk record, forecast, and assessment by experts removes any responsibility from the experts themselves. Thus, the knowledge that underlies the source (ἀρχή) of politics is classified by the possibility of its use in management practice, which significantly narrows the scope of knowledge by reducing it to expertise. As knowledge can give impetus to opposite-direction actions, it reduces the level of rationalization of management. Expert management is limited by unintended consequences, which are seen as inevitable limitations of the system of control. This means that decision-making is no longer a solution to the problem. Since actions of experts increase uncertainty, the rejection of the decision becomes a marginal administrative action that can be accepted in the conditions of a growing danger.

Modern revision of the beginning of politics demonstrates the uncertainty of political power. The existence of parallel processes in the modern management system fully reflects the concept of parapolitics, which is based on Aristotle's doctrine of the principles of separation of power and redistribution. Rancière writes:

"Parapolitics is, first, this centering of political thought on the place and mode of allocation of the arkhaï by which a regime defines itself in exercising a certain kurion" (Rancière, 1999, p. 72).

Aristotle's source of politics no longer opens with philosophical knowledge. The source of politics is power that determines the essence of the politics. In this case, the concept of *ἀρχή* means 'command, order' and refers to dominance. However, Aristotle's works discuss the modes of existence of public power, the essence of which is manifested in the process of communication.

In the concept of parapolitcs, the Greek prefix para-, meaning, firstly, a retreat, a deviation, and secondly, rework, change, indicates there are parallel processes in the structures of power. The public form of power is only a facade, behind which there is an invisible, half-shadow power focused on the redistribution of resources and based on unwritten rules and regulations, rather than law. Public and half-shadow actions of those in power lead in opposite directions. Noteworthy is the fact that synonymous with the prefix para- are such prefixes as trans-, in-, against-, and out-. The source of such politics is completely non-transparent, invisible to society. Half-shadow power follows the principle of one hand washing the other, being a closed circle where there is nepotism, cronyism, protectionism, favoritism, etc. Solovyev writes:

"While in the public sphere the main players are important government officials, political leaders, public opinion leaders, public administrators (of parties and NGOs), political journalists, in half-shadow power their places are occupied by apparatus 'figures of influence' and informal, reference (kinship, family, friendship, community, and so on) coalitions and groupings" (Solovyev, 2011, p. 80).

Reduction of public politics and constant expansion of half-shadow power are accompanied by a shift of boundaries between legal and illegal activities of those in power. Increased illegality reveals itself in the emergence of corruption schemes in

public management practices (bribes, kickbacks). All that was illegal acquired the status of unwritten norms and rules typical of shadow forms of management, including splitting the pot, administrative raiding, illegal channelling of public funds offshore, etc.

It is as if parapolitics demonstrates that the principles and priorities that are proclaimed by those in power can be completely opposite of the policy conducted by the state. The existence of shadow and half-shadow power increases the uncertainty of governance or management, where there is no connection between the words and actions of those who govern and who power belongs to. Discussing any social problems in the framework of public politics turns out to be demagogic 'idle talk', or 'talking shop', which is not supported by any further action. Following its unwritten norms and rules, the government may say one thing and do the opposite, and vice versa, may speak without doing, or do something without speaking.

Closing – in the literal sense of the word – the source of politics allows the non-transparent and uncertain government to act without any explanation or justification of the reasons for its action. The source of such power is determined at the time of an emergency situation. The concept of ultrapolitics is based on K. Schmitt's theory, according to which politics is defined in the 'case of emergency' (extremus necessitatis casus) (Žižek, 2009). In this case, the prefix ultra- indicates the highest characteristic of politics, the function of which 'the state of emergency' performs. To declare 'the state of emergency' is the prerogative of the sovereign, who is not distinguished in a normal condition of peace, safety, and order, but arises in the exceptional case of urgency, when someone with ambitions to assume unlimited power attempts to stand out (Schmitt, 2005). The sovereign creates and guarantees the way in which situation as a whole transpires in its totality, that is, the sovereign confers on himself a higher right–that is independent of the law and does not derive from anything but itself–to make the final decision, which stands apart from the rule of law and proves that the sovereign, in order to create this privileged right for himself, does not need to have the right or precedent to justify it. Deciding, as specifically the legal action of the sovereign, is outside the law, and is therefore unrestricted in legal terms, i.e. becomes absolute. This shows that when the source of politics is closed, the sovereign decision is the only form of public presentation of power.

In an emergency situation the sovereign decision defines friend and enemy. The existential meaning of friend and enemy is always that the enemy is a stranger on the current existential level (Schmitt, 2007, p. 27). Ultrapolitics is an armed struggle against the enemy, the ultimate expression of which is war. Schmitt writes:

"The friend, enemy, and combat concepts receive their real meaning precisely because they refer to the real possibility of physical killing. War follows from enmity. War is the existential negation of the enemy" (Schmitt, 2007, p. 33).

War becomes the only possible action of the government conducting the policy of 'the state of emergency'. The concept of ultrapolitics fulfils the right to engage in war, which is no longer a protection or resistance practice, but is rather a preventive action of the government, justifying their existence through the declaration of 'the state of emergency'. The action of the government toward any citizen, including killing, becomes a preventive measure and remains unpunished.

Thus, it is now possible to observe the transformation of the Greek interpretation of the source of politics. Moving away from knowledge turns the source of politics into a non-transparent source of power, when those who have it begin to act as if they know nothing about the consequences of their actions. This creates an impression that the source of politics is fragmented, anonymous forces represented by countless expert communities, expanding the matrix of bureaucracy, managers, and government administrators. Anonymity de-humanizes (or depersonalizes) power and generates a general feeling of spontaneity of all that is happening with no one in charge. The only action, publicly demonstrating the authority in power and representing the beginning of politics, is 'the emergency situation' caused by social unrest and a threat to society.

Another mode of the beginning of politics is archipolitics, which shows that rejection of knowledge (λογος) as the source (ἀρχή) of politics, returns us to the archaic heritage of the past. The existence of the past in the politics of the present is most fully expressed in Eco's metaphor that the 'Middle Ages have already begun'. More recently, U. Eco argues, our era has been spoken about as the new Middle Ages. The beginning of the new Middle Ages is reflected in the disappearance of state power, decay of states, growing tribal, ethnic, and religious conflicts, the emergence of international criminal mafias, the increasing number of refugees, and the un-

limited proliferation of nuclear and other weapons of mass destruction. Following the medieval model means the onset of political crisis in the future. The state will break into several autonomous subsystems that will be independent of the central government, with hired troops and autonomous courts. Overcoming the crisis will be easier for inhabitants of undeveloped areas, prepared for life and competition in primitive conditions; there will be mass migrations that will lead to the merging and mixing of races and the emergence and dissemination of new ideologies. When the power of law is no longer recognized, and all documents are destroyed, property will be based only on 'the right of custom'. On the other hand, as a result of the imminent decline, cities will consist of both debris and homes suitable for living, where those who will be able to assert (and defend) a claim over them will dwell, and small-scale local authorities will be able to save at least some power only by building fortress walls and fortifications. Then we will find ourselves in a completely feudal system; alliances between local authorities will be based on compromise rather than law; relationships between individuals will be based on aggression and unions built on friendship or common interests; primitive traditions of hospitality will revive (Eco, 1989). The beginning of politics turns out to be its end, but at the end of politics there is nothing but the word or language, able to stop an act of direct action, that will have potential. It is language alone that allows us to avoid such drastic actions as war, which force humanity to face the problem of survival.

The concept of *transpolitics* contains an attempt to overcome the non-transparency of the source of politics. The emergence of the concept is due to the transition of social reality to a new state, when the decisive role is played by the media, the development of which suggests the appearance of an information society (Baudrillard, 1983). If we look at the etymology of the word transpolitics, we will see that its meaning is determined by the Latin prefix trans-, which is used in compound words, and at the same time means over/through/on the other side/behind. In this case, going beyond means the transition of politics into the sphere of information production. Distribution of objective, reliable, and unbiased information is believed to guarantee that the source of politics will be open and understandable.

However, we observe an opposite effect. In transpolitics, the only entity power belongs to is the media, which monopolize the right to represent politics. Politics is gradually absorbed by something that produces it, i.e. the media. Transpolitics removes the difference between reality 'as it really is' and the media's production of political reality. Producing information becomes an autonomous process determined solely by the internal needs of the media system. As a result, a new form of power emerges, the essence of which is expressed in the concept of 'mediacracy'. In mediacracy, power is all that is said in the media, and all that is said should be perceived as something that 'really is'. The media create a stable belief that everyone knows what is happening, but no one acts or wants to do anything.

News production, according to Baudrillard, turns politics into a simulacrum, i.e. an illusion of objective reality. In the media, it can be clearly seen how politics is transforming into a performance that creates a special form of 'spectacular policy', which opens the space of visibility (Debord, 1994). In the space of visibility, the performance establishes a visual procedure that involves the public in watching media representations. The greater the viewer involvement is, the more spectacular politics becomes, following the model of entertainment talk shows, court interviews, or TV debates featuring leaders of parliamentary parties. In the information flow, politics is no longer seen as a logical and consistent process and seems an excess. Politics becomes a subject and source of information only in the case of corruption revelations, mistress scandals involving government officials, violations of legal, moral, political correctness norms, etc.

In 'spectacular politics', the word is replaced by a visual image that is mixed with other forms of representation. The image turns out to be a distorted language, altering the linguistic nature of human beings. The words of a language are re-denoted and lose traditional meanings and connotations. The media, claiming full power, generate a new language where meanings of such significant concepts as 'life' and 'death', 'war' and 'peace' are replaced with euphemisms. Contemporary wars, on the surface of meaning, manifest themselves as a movement for justice and humanism, i.e. pretend to be a kind of 'humanitarian mission' or 'pacifist humanitarian intervention' aimed at promoting peace. In this case, war is presented as 'protection of peace', where any attack is only a 'strategic alternative to defence' or a preventive action aimed at preserving security in the world. In turn, the state of the world is regarded as an endless series of 'states of emergency', gradually leading the world to a state of 'permanent war'.

The language of the media loses its main function – the ability to express thoughts. The image is always presented in only one way, i.e. the way it is presented in the media. In this sense, the image does not carry anything imaginary; it literally means 'what is seen'. The 'image' no longer refers to anything else but itself as a kind of visual picture, the meaning of which is determined performatively, in the process of commenting. The comment, as an explanatory note, attaches additional meaning to the visual image, and this forms the 'secondary system of meanings'. This 'secondary system of meanings' is a myth, in which the primary sign of the semiotic system is the visual image that combines the primary signifier and signified to becomes a signifier itself. The unity of the secondary, mythological signifier (form) and the corresponding secondary signified (concept) produces the sign of the second order – the meaning (Barthes, 1972). One can say that at the limit of linguistic expression, information is presented in the form of a myth, which unfolds in narratives/stories about events that may have occurred or may have never occurred. The media are, in essence, a means of producing political myths, which exist only in the media and are supported only by the media. Such political myths are necessary to maintain the existence of the media themselves.

Society loses the ability to distinguish between the media illusion and the political reality. The disappearance of boundaries between illusion and reality causes 'trance' as the limit state, which occurs when the distance between illusion and reality is reduced to 'zero'. In transpolitics, the prefix trans-, which was originally interpreted as moving/transition into a different reality, acquires a different meaning, indicating the state of society which is characterized by a clouded state, similar to insanity or madness, or hypnotic trance, to which the media condemn 'the society of the spectacle'. Being in a 'trance', society loses common sense. One can say that transpolitics extends the idea of the end of politics, allowing it to last endlessly in stories and narratives that are produced by the media, serving an endless source of political mythology, mainly focused on entertaining the public.

Between the indefinite beginning and the indefinite end of politics, there is an infinite space of the politics of life, or biopolitics. The etymology of the word 'biopolitics' is determined by the prefix bio-, which is used in compound words and is derived from the Greek βίος, denoting life, and the Latin bios-, corresponding to the meaning of the word 'biological'. The prefix bio-, providing the basic idea of the word, is a combination of the Greek and Latin interpretations in the way that biopolitics expands the notion of social life to the biological/natural level of society's existence. In this case, life is understood as zoe, which is implemented at the level of species (i.e. humankind). Modern biopolitics abolishes the semantic difference, introduced by Greek philosophy, between life as zoe, meaning the fact of life common to all living beings, and βίος, pointing to the political life of society. In Greek tradition, the political life of a polis βίος πολιτικός was separated from natural, or biological, existence, which focused on the household as οἰκία, where there was βίος ἀπολαυστικός – life for the sake of pleasure. The modern concept of biopolitics imposes a limit on the Greek understanding of politics as βίος πολιτικός.

Biopolitics begins when political life (βίος) is interpreted simply as natural/biological life (zoe), which implicates life itself within the sphere of government control and supervision. Biopolitics involves the generic principle of life, which is included in politics in the state of 'bare life' (Agamben, 1998). All biopolitics is presented as a way to control the processes of life and expressed in supporting the existence of society. The practice of managing the life of society develops in two main directions – regulation of population life and continuation of individual life (Foucault, 2013). Organization of life is centred around the human race, which serves as the basis for control of the biological processes of reproduction, birth, and mortality, health status, life expectancy and longevity. Regulatory methods of control over the human race are realized through population biopolitics. In this case, decreasing child mortality and increasing life expectancy become an indicator of care taken by the government of the population. The government is responsible for the survival and growth of the population as a kind of biomass. The government takes cares of the population only to maintain itself as a kind of biopower, regulating the life of the social body from inside.

The life of an individual is controlled by the anatomo-politics of the human body, including the individual as one of the species of living beings that functions primarily as a kind of a living/biological body. Anatomo-politics of the human body, according to M. Foucault, is focused on disciplinary control over the life of the individual's body. Disciplinary techniques of control and persuasion develop

the life of the individual's body in accordance with the automatism of habit. The requirement of disciplinary power to live in a certain way reaches the unconscious/non-conscious level of the individual's life as the level of ultimate subjectivity. The more obedient the body becomes, the more effective the life of the individual is and the more useful such a body is for the government. The obedient body of the individual allows the government to act through itself and in itself. The individual's body is marked with the social content that the government assigns it. Any claim of individuals to their bodies leads to a change in government strategies. In biopolitics, disciplinary strategies of suppressing the individual's body are replaced by the strategies of stimulating the life of the body.

The individual becomes her own investor who assumes the costs of improving the life of the body. The strategy of stimulating the body's life does not require any additional investments from the government. The social efficiency of the individual begins to be measured through both innate and acquired abilities. Innate abilities are expressed through the natural abilities and inclinations of the individual, which are developed by improving both the external physical parameters of the body (gymnastics, muscle development) and the internal ones, thus changing the genetic inheritance of a population over time. The strategies of the control-stimulation of the individual's body are directed toward the very core of life, i.e. they infiltrate the life process reaching up to its elementary component - the human genome. Modern bio-genetic experiments in the form of genetic engineering, the use of GMOs, cloning, transplanting and growing artificial organs, and artificial immortality are aimed at improving the individual's physical and mental abilities. Biogenetics seeks to free life from biological limitations, which would drastically increase life expectancy, and the bio-genetic makeup of human beings ceases to be regarded as a result of natural evolution.

Natural biological process management is embedded in the natural order of life as its necessary component, the existence of which is determined by the processes of life itself. In biopolitics, the administrative apparatus is a de-politicized element of life. One can say that biopolitics diffuses throughout the technology of life and reaches its limit in the everyday politics of individuals who are concerned with increasing life-supporting resources. At the level of life technology, each individual turns out to be one's own manager, a rational automatic machine that regulates the processes of one's own life. The emergence of 'everyday politics' indicates the absence of a grand 'Politics' (starting with the capital letter) (Bauman, 2001). Management of life, deprived of the political beginning, is the essence of post-politics, where the prefix post- indicates the overcoming of ideological interpretations of politics and the transition to a politics of expert management and administration (Žižek, 2009).

Noteworthy is the fact that in socio-philosophical discourse the concept of "politics" is gradually pushed out and replaced with police, representing a version of the interpretation of management practices and its technological components. Etymologically, the word police – polizei – is as close as possible to 'policy'. Currently, the negative meaning, which the concept of police as a punitive supervisory authority of coercion and direct suppression traditionally had, is disappearing. In the past, police powers were mostly limited to the apprehension of offenders and supervision of public order, while now the term 'police' is becoming synonymous with the overall governance and control over society. This suggests a return to the original, positive meaning of police (as relating to policy), indicating a change in political practice that entails the need to identify new meanings of the words used to describe reality itself.

The positive meaning of 'police' is restored through reference to the original meaning of the word. Exploring the origins of police M. Foucault points out that, firstly, "police" was used to refer to a community that was governed by public authority, secondly, it was used to denote a set of acts relating to the management of this community, and thirdly, in the 17th century, the word 'police' began to be understood as a set of tools to facilitate the growth of the state's forces. Foucault writes:

"The objective of police is therefore control of and responsibility for men's activity insofar as this activity constitutes a differential element in the development of the state's forces" (Foucault, 2007, p. 417).

Police as a management practice includes both public and private mechanisms to ensure public tranquillity and the security of the state. Police practice becomes an immanent principle of society's organization, applicable to all levels of the population's existence: health, education, work, etc.

Erasing the semantic differences between 'policy', performing positive tasks, and 'polizei', performing negative tasks, leads back to the phenome-

non of opposites changing places and 'overturning' one another. As a result, 'police' gets a positive meaning of practical activity while 'politics' acquires a purely negative one. 'Policeman' became a common name for all who are involved in the normal practice of supporting the life of society: a government member, an expert, a lawyer, a doctor, a teacher, a sociologist, because public opinion polls have become a form of police control. In police regimes, such actions as protests, strikes, and demonstrations are opposition actions that disrupt public tranquillity and the security of the state since they introduce confusion and disorder. These actions are classified as legal issues, and therefore, must be suppressed. In a police regime, politics has no source, i.e. is anarchic in nature (Rancière, 1999). In the state of anarchy, politics can be implemented anytime and anywhere, but its origin is connected to the violation of the established social order. There is complete nihilisation of politics, which is set to the 'zero' level of existence, where there is a kind of politics because the state continues to operate, the government does something, and people live, yet at the same time, there is no politics.

In the structures of social and philosophical discourse prefixing in conceptualizing politics reaches the limit of linguistic possibilities. The limit can be regarded as a 'turning point' where the state of *post*-politics acquires the meaning of *meta-*. Meta- is one of the most productive prefixes in conceptualizing politics because it creates the possibility of transitioning to the metaphysical foundations of politics. The concept of 'metaphysics' derived from the Greek 'μετα τα φυσικα' refers to the way in which a study carries political 'meta' issues beyond questions of power and ideology. The plurality of meanings of the language prefix meta- are reduced to two meanings – 'after, behind' and 'among, together, accompanied by'. The first meaning of meta- refers to a situation of transition, which is denoted by the Latin prefix 'post-'. However, the idea of movement is found in the Greek word μεταβωλή, denoting turnover, transformation. The second meaning of 'meta' is communicated with the Latin prefix 'trans-'. Heidegger writes:

"The meaning of changeover, of 'turning away from one matter toward another', of 'going from one over to another', came out of a purely positional meaning" (Heidegger, 1995, p. 39).

Going beyond (post-) the prefixal ways of interpreting the noun 'politics', which has completely exhausted itself, turns (trans-) to the ontological foundation of political life as a system of philosophical knowledge.

Appeal to ontological grounds returns philosophy to itself as a source of apolitical existence. In this case, a- is a negative prefix indicating the absence of the characteristic or property expressed by the main part of the word. The prefix a- is productive only to form adjectives and absolutely unproductive to form nouns. While literally or grammatically the prefix a- refers to negation, semantically it makes a distinction between what is called the reality of politics and political reality, i.e. between the being of politics related to power and ideology, and political existence. As much as the difference between the reality of politics and political reality reveals itself through the negative prefix a-, such a distinction is, in fact, an ontological way of distinguishing between being and existence.

In this case, Badiou's concept of metapolitics is becoming especially important (Badiou, 2005). However, Badiou uses the prefix meta- to denote the transition to a higher level of understanding, which he associates with the logical analysis of politics, opening the level of meta-philosophical reflection. Philosophical meta-reflection is understood as part of politics since it undermines 'the established intellectual order' as a certain regime of truths accepted in political philosophy, legitimizing a certain type of politics.

In turn, the ontological 'turn' takes us to the metaphysical grounds of political reality. In this case, the prefix meta- means the transition from the noun 'politics' to the adjective 'political', which is denoted by the Greek word πολῖτεια, returning us to the philosophical roots of political life as a knowledge system based on the original identity of being and thinking. In philosophical discourse oriented at the study of political systems of knowledge, the grammatical structure is changing. There prefixing as a way of interpreting is replaced with suffixing, with suffixes being used to form vague adjectives that denote only a sign, a property, or a quality of reality.

In Greek philosophy, an apolitical source of political life as a system of knowledge was thinking (λογος) and action (πραξις). Thinking is action since 'the thought thinks'. As much as thinking thinks, the activity of thinking is aimed at making activity products, i.e. is the action of producing, within which the thought is thought through thinking. The product of the activity of thinking is the idea, the essence or the meaning of which is expressed by

language. Thinking that exists in the identity with language is a productive action (Heidegger, 1968). In turn, action as πρᾶξις is realized through the word, which loses the linguistic meaning λέξις and is understood as λογος, i.e. it represents a conscious linguistic activity that exists in the identity of language and thought. The notion that any idea asserts a justified claim to the status of dominant ideology is dismissed with philosophical reflection, which, on the one hand, helps to illuminate the degree of the presence and intervention of the ideology in social life, and on the other hand, preserves thinking and language as the source of political existence..

So, to summarize, it can be stated that the trend of prefixing reflects the desire of researchers to break the traditional link between politics and ideology and to justify a concept of politics that would be more in line with the realities of the management state. As a result, researchers revise the basic concepts of the beginning, the end, and the process of political life. Moving away from philosophy as a necessary source of politics is one of the ways to overcome ideology in politics. Post-ideological politics gives form to the idea of police and police management, and this idea is more total(izing) than any previous ideology, since it diffuses throughout the technology of life and permeates all levels of social existence, including the natural/biological one, which has never before been infiltrated by politics. The idea of police management is implemented under the banner of concern for the safety of life, which is possible only if social life as a whole and the life of each individual is absolutely transparent and controllable. The transparent society susceptible to absolute control becomes the only idea to be implemented–not in the future, but in the present. We believe that the ontological 'turn' returns us to the philosophical source of politics. Philosophy demonstrates the conventionality of the beginning of politics and the absoluteness of thinking of λογος as the beginning of politics, restoring the original value of action πραξις.

References

Agamben G (1998) *Homo Sacer: Sovereign Power and Bare Life*, trans. Daniel Heller-Roazen. Stanford: Stanford University Press.

Agamben G (2013) *Qu'est-ce que le commandement?*, traduit de l'italien par Joël Gayraud. Bibliothèque Rivages.

Badiou A (2005) *Metapolitics*, trans. Jason Barker. New York: Verso.

Barthes R (1972) *Mythologies*, trans. Annette Lavers. New York: The Noonday Press.

Baudrillard J (1983) *In the Shadow of the Silent Majorities or, The End of the Social and Other Essays*, trans. Paul Foss, Paul Patton and John Johnston. New York: Semiotext(e).

Bauman Z (2001) *The Individualized Society*. Cambridge UK: Cambridge University Press.

Debord G (1994) *The Society of the Spectacle*, trans. Donald Nicholson-Smith. New York: Zone Books.

Eco U (1989) *The Middle Ages of James Joyce: The Aesthetics of Chaosmos*, trans. Ellen Esrock. Cambridge, Mass: Harvard University of Chicago Press.

Foucault M (1980) *Truth and Power* // Foucault M. Power/Knowledge. Selected Interviews and Other Writings 1972–1977, ed Colin Gordon and trans. Colin Gordon, Leo Marshall, John Mepham, Kate Soper. New York: Pantheon Books.

Foucault M (2007) *Security, Territory, Population: Lectures at the Collège de France, 1977–78*, ed. Michel Senellart, trans. Graham Burchell. New York: Palgrave Macmillan.

Foucault M (2013) *Lectures on the Will to Know: Lectures at the Collège de France, 1970–1971*, and Oedipal Knowledge, trans. Graham Burchell, ed. Daniel Defert. New York: Palgrave Macmillan.

Foucault M (2008) *The Birth of Bio-Politics: Lecture at the College de France 1978–79*, trans. Graham Burchell, ed. Arnold I. Davidson. Basingstoke: Palgrave Macmillan.

Heidegger M (1995) *The Fundamental Concepts of Metaphysics: World, Finitude, Solitude*, ed. John Sallis, trans. William McNeill and Nicholas Walker. Bloomington, Indiana: Indiana University Press.

Heidegger M (1968) *What is Called Thinking?*, trans. J. Gray. New York: Harper & Row.

Rancière J (1999) *Disagreement and Philosophy Politics*, trans. Julie Rose. Minneapolis – London: The University of Minnesota.

Schmitt C (2005) *Political Theology: Four Chapters on the Concept of Sovereignty*, trans. George Schwab. Chicago: University of Chicago Press.

Schmitt C (2007) *The Concept of the Political. Expanded Edition* (1932), trans. George Schwab. Chicago – London: University of Chicago Press.

Soloviev Alexander (2011) Latent structure of governance, or the play of shadows on the face of authorities. *Polis. Political studies* 5: 70–98.

Wallerstein I M (1995) *After Liberalism*. New York: New Press.

Wallerstein I M (1999) *The End of the World As We Know It: Social Science for the Twenty-first Century*. Minneapolis: University of Minnesota Press.

Žižek S (2009) *The Ticklish Subject: the Absent Centre of Political Ontology*. London-New York: Verso.

Discussion

LIBERAL DEMOCRACY AND CULTURAL DIVERSITY – BETWEEN NORMS AND FACTS

Plamen Makariev
Sofia University "St. Kl. Ohridski"
Sofia, Bulgaria
e-mail: makariev@abv.bg

Abstract:
This article has been written in response to the texts by Richard Robson ("In What Sense is Multiculturalism a Form of Communitarianism"), and Slobodan Divjak ("Communitarianism, Multiculturalism and Liberalism") with which the Balkan Journal of Philosophy (vol. 10, no 2, 2018) started a discussion on the theme Liberal Democracy and Cultural Diversity. I try to contest the position of these two authors–that multiculturalism and communitarianism belong to one and the same paradigm in political philosophy–by pointing out essential liberal normative elements in multiculturalist theory. My main thesis is that in order to clarify the relation between multiculturalism and communitarianism, we have to differentiate between descriptive and normative communitarianism. The latter is guided, in my opinion, by values, which stand in stark contrast with the liberal ones, whilst this is not the case with multiculturalism.

Key words: multiculturalism, communitarianism, liberalism, universalism, particularism.

Introduction

My main point of reference for this study is the unquestionably humanist nature of the values and norms which characterize liberal democracy. The public policies concerning cultural diversity in such societies cannot involve discrimination, humiliation, or oppression of any of their citizens. On the other hand, any society–be it liberal-democratic or not–cannot afford a disintegration of the social bonds that maintain the self-consistency of its functioning, cannot afford a shattering of social peace, etc. How can these two determinants of social life–justice and order– be harmonized? Quite a few answers to this question have been proposed from the perspectives of different paradigms of social study, and I welcome the idea of the *Balkan Journal of Philosophy*'s editors to start a discussion on these issues by publishing in Issue 2 of Volume 10 (2018) of the Journal a short text by Richard Robson, titled "In What Sense is Multiculturalism a Form of Communitarianism", and Slobodan Divjak's article "Communitarianism, Multiculturalism and Liberalism". In what follows I will try to contest their position–that multiculturalism and communitarianism belong to one and the same paradigm in political philosophy–by pointing out essential liberal normative elements in multiculturalist theory.

First I'll present in brief some of the social and political problems which ensue from the "tension" between cultural diversity and the liberal-democratic social order. Then I'll outline the main philosophical approaches to them, trying to "diagnose" the conceptual flaws, which prevent these theories from serving as effective methodologies for public policies that would deal more or less successfully with the problems in question. And finally, I'll present my "best guess" concerning the prospects of the possible developments in this field.

The political issues

Judging from the facts of social and political life in 'Western' societies, as well as from the predominant interpretations of these facts in the mass-media and in social studies, I draw the conclusion that the problems which result from the manifestation of racial, ethnic, and religious differences in liberal-democratic social environment can be ordered on an imagined "axis" between two "poles": a universalist and a particularist one.

In societies where public policies follow universalist strategies, i.e. the importance of the differences just mentioned is relativized in various ways, this results in exerting an assimilationist pressure on the members of minority communities. Policies of such nature may be, first of all, unjust if they do not leave these people the right to choose freely whether to assimilate into the mainstream culture or not. In many cases such policies bring about worldview disorientation and moral degradation among these citizens; they may suffer from loss of identity or loss of the meaning of life, which may lead ultimately to anomy, alcoholism, drug addiction, criminality, etc.

On the other "pole", the particularist one, the importance of minority identities is exaggerated and diversity is "celebrated". This was the case with multiculturalist policies in North America and in Western Europe throughout the last decade of the XXth century and the first decade of the XXIst. As is well known, these policies brought about (not everywhere, but in a number of countries) a self-segregation of minority communities, formation of immigrant "enclaves", and development of "malignant" subcultures that nourished religious–especially Islamic–fundamentalism and radicalism. In short – the particularist approach to cultural diversity endangered the unity of these societies.

What to do?

The philosophies which address these issues can be considered in such a frame of reference, too. The extremely universalist paradigm in this regard is "orthodox" liberalism, which aims ultimately at the "erasure" of cultural diversity within the respective societies. The extremely particularist position is represented by *normative* communitarianism, which provides theoretical justification of racist, nationalist, and totalitarian ideologies, that affirm the supremacy of one racial, national, ethnic, or religious identity above all others. And at various points on the "axis" between the two "poles" we may find political liberalism in various versions[1], liberal multiculturalism[2], "classical" multiculturalism[3], certain theories of deliberative democracy[4], etc.

How do the authors, who adopt one or another of these approaches to cultural diversity justify their choice, i.e. how do they explain why precisely the combination of universalist and particularist policy elements, which they recommend, is the optimal one? Why do they think that only in this way, and not otherwise, the right compromise between satisfying the cultural needs of minority communities and preserving the integrity of society can be found? What bothers me in this respect is that quite a few philosophers and political scientists rely in this regard on *introspection*.[5] A scholar would decide what the right choice between loyalty to universal humanistic norms and to one's own community ought to be by judging from her/his own personal experience – from what s/he *feels* in certain situations.

For example, from a certain perspective, the liberal "unencumbered self" may look like someone without true identity, someone unable to really love and care about other persons in their capacity as living human beings, rather than abstract individuals. "To imagine a person incapable of constitutive attachments such as these is not to conceive an ideally free and rational agent, but to imagine a person wholly without character, without moral depth. For to have character is to know that I move in a history I neither summon nor command, which carries consequences nonetheless for my choices and conduct. It draws me closer to some and more distant from

[1] John Rawls, *Political Liberalism* (New York, Columbia University Press, 1993); John Rawls, "The Idea of Public Reason Revisited," *The University of Chicago Law Review*, 64, no 3 (1997): 765–807.

[2] Will Kymlicka, *Multicultural Citizenship. A Liberal Theory of Minority Rights* (Oxford, Clarendon Press, 1995)

[3] Charles Taylor, "The Politics of Recognition". In *Multiculturalism. Examining the Politics of Recognition*, edited by Amy Gutmann, 25–74. (Princeton N.J: Princeton University Press, 1994)

[4] Michael R. James, *Deliberative Democracy and the Plural Polity.* (Lawrence, KA, University Press of Kansas, 2004); Jorge Valadez, *Deliberative Democracy, Political Legitimacy and Self-Determination in Multicultural Societies* (Oxford, Westview Press, 2001); Monique Deveaux, *Gender and Justice in Multicultural Liberal State* (Oxford, Oxford University Press, 2006)

[5] I am using the psychological term *introspection*, instead of the philosophical *self-reflection*, because the problem here is, in my opinion, that the authors in question try to substantiate normative claims by referring to a factual state of affairs.

others; it makes some aims more appropriate, others less so."[6]

From the opposite perspective, there is no such thing as a self-consistent cultural identity at all, i.e. there is nothing that deserves to be valued as such by the people who belong to these groups, such that they should be morally obliged to care about it, to be loyal to their community, etc. Could we speak of a particular identity of a person who "…may live in San Francisco and be of Irish ancestry, […] learns Spanish, eats Chinese, wears clothes made in Korea, listens to arias by Verdi sung by a Maori princess on Japanese equipment, follows Ukrainian politics, and practices Buddhist meditation techniques".[7] Where does that leave the internal consistency of behavior, the distinct character that accounts for cultural identity? Isn't the latter an idealization which is too strong for postmodern times and which should not be taken seriously or, even less so, considered to be a factor in defining minority policies?

Often we come across such manners of thinking, which actually derive norms from facts. Someone feels in a certain way concerning the relation between universal values and norms, on the one hand, and cultural identities, on the other, and draws from this the conclusion that everyone else *should* feel the same. If someone doesn't, s/he is wrong. From a particularist point of view the universalist attitudes are morally incorrect, because they do not correspond to human nature, which *does need* identity. On the contrary, if one thinks in a universalist way, s/he believes that the particularist mentality is anachronistic – a manifestation of tribal idiosyncrasies–and it is incompatible with a humanist morality. So, who is right and who is wrong?

It is an undeniable fact that there are people in the world who believe in particularist values, and there are also people who subscribe to universalist ones. Do we have the right to derive moral norms from these facts? Can we morally justify particularist or universalist minority policies by referring to the fact that some of the people concerned nourish sincerely held beliefs of this or that kind? In my opinion, this is a methodological challenge, which should not be ignored,when we comment on the debates about cultural diversity and liberal democracy.

The universalist and the particularist paradigms

Assimilation

Concerning minority policies, one obvious possibility is the exertion of assimilation pressure on members of such communities. This may be done openly, through public legitimization of policies of this kind. However, it may also be pursued discreetly, where minority members are subjected to informal pressure aimed at motivating them to seek to improve their quality of life and prospects for professional and social advancement by disowning their identity.

Two quite different models of assimilation are possible in theory and found in practice. One is characteristic of the so-called "civic" nations. In these nations assimilation is conducted not by replacing the minority identity with another, but by its systematic belittling and encouragement of its subjects to identify with the nation as a civic, rather than cultural, community. Individuals are expected to be guided, above all, by universal values instead of cultural traditions and group solidarity. Emblematic in this respect is the French ("republican") model of dealing with minority issues – at least in its traditional form[8].

Nominally, this kind of pressure on cultural identities is not qualified as assimilationist because it does not aim at changing them, but rather at relativizing their importance. In fact, however, what we have here is the clandestine assimilation of minorities by the majority on the pretext of securing national civic unity.[9] Majority is equated with normalcy[10], national identity is equated with the cultural identity of the dominant group because national

[6] Michael J. Sandel, "The Procedural Republic and the Unencumbered Self," *Political Theory*, 12, no 1 (1984): 81-96; 90

[7] Jeremy Waldron, "Minority Cultures and the Cosmopolitan Alternative", in *The Rights of Minority Cultures*, edited by Will Kymlicka, 93 – 122. (Oxford, Oxford University Press, 1995) 95.

[8] See for example Jonathan Laurence, "The New French Minority Politics." The Brookings Institution (March 2003). http://www.brookings.edu/fp/cusf/analysis/laurence.pdf, accessed April 28, 2017; Adel Iskandar and Hakem Rustom. "From Paris to Cairo. Resistance of the Unacculturated." *The Ambassadors* 9, no. 1 (2006). http://www.ambassadors.net, accessed December 2nd 2017.

[9] At the cost of minimizing the significance of – allegedly – *all* cultural identities in public life.

[10] See Tibor Várady, "On the Chances of Ethnocultural Justice in East Central Europe." in *Can Liberal Pluralism Be Exported? Western Political Theory and Ethnic Relations in Eastern Europe*, edited by Will Kymlicka and Magda Opalski, 135 – 149 (Oxford, Oxford University Press, 2001).

identity cannot be entirely culturally neutral.[11] Even with the best intentions in the world, it is impossible to teach history, literature, and the national language at school in a culturally neutral way (in the rare cases of official bilingualism, the two languages are taught in a culturally balanced, rather than neutral way). The very idea of a *culturally neutral national culture* is an oxymoron, therefore, it is impossible to realize in practice.

In this respect I disagree with S. Divjak, whose article "Communitarianism, Multiculturalism and Liberalism", published in the Balkan Journal of Philosophy (2018/2), started the present discussion on liberal democracy and cultural diversity. There he states: "By official recognition of some groups, the state gives certain privileges to them. Their members have not only civil but also specific cultural and political rights, while the members of majority cultural group have only civil rights. When ethnic groups acquire ethnic autonomy, they can interpret themselves as the people in the ethnic sense, while the other citizens remain the people in a political sense."[12]

This argument is also used quite often by other authors who subscribe to an assimilationist perspective. However, it is at odds with the practice that I mention above to equate national identity with the ethnic identity of the majority. Education, the mass media, national literature, the arts – all of these affirm in one way or another the value of the majority culture, the artistic achievements of representatives of this community, heroic exploits of its members in the past, role models that are characteristic of its culture, etc. How can we ignore all this and agree that the national identity, even in the most liberal societies, is culturally neutral – that it is of an entirely civic nature?

By presenting these considerations I do not claim that assimilation is something bad in itself. If it is practiced in ways that provide members of minority groups the option to voluntarily identify with the majority culture (or not)– freely and for whatever reasons– there should be nothing morally wrong about it from a liberal-democratic perspective. The opposite, i.e. a denial of such an opportunity, would be a kind of segregation. What I am arguing against here is the exertion of assimilationist *pressure* on such people, and the hypocritical claims that this is not forced or coerced assimilation, but integration into the civic structure of society.

The other mechanism of assimilation is applied in the so-called "ethnic" nations, where members of minority groups are pressured to identify with the majority. A glaring example of this type of assimilation was the so-called "Revival Process" in Bulgaria.[13] To motivate the Bulgarian citizens of Turkish ethnicity to change their identity, the concept was developed that all members of this group were distant descendants of ethnic Bulgarians who, during Ottoman rule, had been Islamized and eventually assimilated ethnically, as well. It was claimed, therefore, that a process of "re-Bulgarianization" would not constitute assimilation, but rather a return to the authentic, true identity of these people – hence the official name of the campaign, "Revival Process":

> *The main argument was historical. It included the thesis that all Bulgarian Turks were descendents of Bulgarians converted to Islam during Turkish rule rather than of Turkic colonizers; that they were now reawakening from a centuries-long slumber; that they were now happy to restore their "true names" along with their true national self-identity.*[14]

Actually, this type of assimilation–the "direct" kind, so to speak–is not universalist, but rather particularist in nature. The values that are used as a reference point to justify such policies are not civic, but culturally specific. This brings us to the domain of the particularist methodologies of social activities.

[11] See e.g. Charles Taylor, "The Politics of Recognition", in *Multiculturalism. Examining the Politics of Recognition*, edited by Amy Gutmann, 25 – 74. (Princeton N.J, Princeton University Press, 1994). 43

[12] Slobodan Divjak, "Communitarianism, Multiculturalism and Liberalism", *Balkan Journal of Philosophy*, 10, no 2 (2018): 147–163, 158

[13] A large-scale campaign for the assimilation of ethnic Turks in Bulgaria launched in the winter of 1984/85 which included, in the first place, motivating – through various forms of pressure – these Bulgarian citizens to apply to the authorities with a request that their names be changed (from Muslim to Christian or Slavic ones). However, since the ideological justification for this campaign was based on the thesis that Bulgarian Turks were, *in essence*, descendants of long-since assimilated ethnic Bulgarians, the campaign also included a ban on speaking Turkish in public, and even *ex-officio* renaming of deceased persons in the civil status records, as well as a number of restrictions on the practice of Islam.

[14] Krassimir Kanev, "Law and Politics on Ethnic and Religious Minorities." in *Communities and Identities in Bulgaria*, edited by Anna Krasteva, 55 – 95. (Ravenna, Longo Editore, 1998), 77

Discrimination

Most generally, discrimination is regarded as *disadvantageous* treatment of people, because of some kind of difference which singles them out in mainstream society. These differences can be of whatever nature – age, sex, sexual orientation, physical disabilities, national origin, political affiliation, etc. The ones that are of interest to us in our case are the cultural differences – race, ethnicity, religion, national affiliation. One of the most influential theories, which offer explanations of this phenomenon nowadays, is the theory of social identity[15]: a person who identifies in some way with an ingroup is always "tempted" to discriminate against members of outgroups.

If we adopt this approach, we cannot but associate discrimination with inter-communal relations, and this brings us to communitarianism as theory, ideology, and social practice. Theoretically, communitarianism is the main opponent of liberalism, insofar as from its perspective particularist values are opposed to universal ones as their alternative. For example, solidarity vs justice, and loyalty to one's community vs freedom. While in a liberal frame of reference the Right (justice) has always been given priority over (what is perceived to be) the (common) Good, the opposite is true in a communitarian one. In the name of the common Good any sacrifices concerning members of the respective community are justified, including violations of their human rights, restrictions of their freedoms, etc. For the same reason, any acts of discrimination against outsiders can seem legitimate in the eyes of the community's members.

However, in order to avoid misunderstandings concerning the relations between communitarianism and multiculturalism, I think that we should differentiate between *descriptive* and *normative* communitarian theories. The most influential philosophical conceptions of this sort, put forward by scholars like M. Walzer, M. Sandel, A. MacIntyre[16] are, in my opinion, of descriptive nature. The authors point out reasons why people tend to identify with communities that they belong to in terms of their ethnicity, religion, or nationality; reasons why their self-understanding is (or should be), first of all, a self-understanding as a member of this or that community, and only after that, as citizens of their state; reasons why solidarity with their fellow members matters (or should matter) for them, and so on.

However, none of these scholars, to the best of my knowledge, would relate these considerations to himself personally. Neither of them puts himself "in the shoes" of a person with a communitarian mindset– i.e. none of them claims that his religion is the only true one, that for him discrimination of outgroup members is acceptable if it is practiced for the good of a community that he identifies with, etc. This is what I mean when I qualify the most well known communitarian conceptions as descriptive. They are characterized by some degree of relativism, insofar as their authors seem to recognize the right of members of different communities to defend 'the common Good' as it appears to each of the respective communities to which individuals belong. Still, this does not constitute normative communitarianism. Examples of really normative communitarian views would be racist ideologies, national socialism, fundamentalist religious doctrines, etc.

The Case of Multiculturalism

As I mentioned in the beginning of this article, the two texts by which the Balkan Journal of Philosophy started the discussion on liberal democracy and cultural diversity are by Richard Robson and Slobodan Divjak. Both authors maintain that there is a lot in common between communitarianism and multiculturalism. The very title of Robson's text says "In What Sense is Multiculturalism a Form of Communitarianism".[17] And S. Divjak points out in his article that both multiculturalism and communitarianism are contextualist theories[18], that they all assert that the community's Good should have priority over the Right, that from a communitarian point of view Reason is not independent from cultural tradition[19].

Without denying the existence of such commonalities between these two kinds of theories, I'll

[15] Henri Tajfel, *The Social Psychology of Minorities*. (London, Minority Rights Group, 1978).

[16] Michael Walzer, *Thick and Thin. Moral Argument at Home and Abroad*, (Notre Dame, University of Notre Dame Press, 1994); Michael J. Sandel, "The Procedural Republic and the Unencumbered Self," *Political Theory*, 12, no 1 (1984): 81-96; Alasdair MacIntyre, *Whose Justice? Which Rationality?* (Notre Dame, University of Notre Dame Press, 1988).

[17] Richard Robson, "In What Sense is Multiculturalism a Form of Communitarianism," *Balkan Journal of Philosophy*, 10, no 2 (2018): 145 - 146

[18] See Slobodan Divjak, "Communitarianism, Multiculturalism and Liberalism", *Balkan Journal of Philosophy*, 10, no 2 (2018): 147–163, 161

[19] See Slobodan Divjak, 150

try to show, that multiculturalism is characterized not to a lesser degree by intercultural tolerance, which is an essentially liberal, universalist norm. The recognition of the value of particular cultures in a multiculturalist vein is necessarily a *reciprocal* act. I recognize the value of your culture, and I expect that you do the same concerning mine – a position which is incompatible with normative communitarianism.

How do multiculturalist authors justify the moral duty to recognize the worth of others' cultures? Here I will briefly comment on two of the best-known arguments, which fit nicely into the context of liberal values. Arguably the most popular one of these is formulated by Charles Taylor in his seminal essay on "The Politics of Recognition"[20]. According to this argument, it is our duty to accept all cultures (or at least those that have proven their significance historically, having formed the mentality of entire societies for long periods of time)[21] as valuable in an absolute sense – that is, regardless of the point of view from which this judgment is being made. At the beginning of his essay, Taylor refers to a more or less obvious characteristic of identity – that it is dialogical in nature, and that we define our identity and answer the question, "Who am I?," based on not only what we see in ourselves, but also on what the persons who are important in our lives see in us.[22] It follows from this that the demonstration of a negative, or even just contemptuous, attitude toward the Other cannot be taken as an innocuous expression of personal opinion, for "a person or group of people can suffer real damage, real distortion, if the people or society around them mirror back to them a confining or demeaning or contemptible picture of themselves"[23]. Hence, recognition is not just "a courtesy we owe people. It is a vital human need"[24].

With this argument, Taylor presents recognition as something that is owed to every individual to the extent that refusal to grant recognition would be *unjust*. In his view, what should be recognized is not one's identity in a general, abstract sense, but precisely the cultural identity of the individual. It turns out that recognizing the worth of another's culture is a matter of *fairness*. That is why the recognition paradigm, as interpreted by Taylor, fits into the context of liberal values. It is one of the sources used to justify, theoretically and morally, multicultural public policies from a liberal point of view. Following the logic outlined in Taylor's essay, the conclusion that the self-respect of an individual depends not only on the way she views herself, but also on the way she "sees" herself through the eyes of others, leads to slogans of the "let's celebrate diversity" type, which were very popular in the 1990s.

The next argument I will look at belongs to another authoritative exponent of multiculturalism, Will Kymlicka. It is based on the ideal of the liberal neutrality of the state. Kymlicka insists that this ideal should be complemented by the recognition of the importance of what he calls "*societal culture*." The Canadian philosopher questions the liberal ideal of the neutrality of public institutions with regard to conceptions of the good life, including culture, and of limiting their powers only to the fair regulation of social interactions. According to Kymlicka, the state *de facto* actively participates in the functioning of a "societal culture" which is (at least partly) institutionalized, as well as laden with cultural content. He proposes the following definition of societal culture:

> *a culture which provides its members with meaningful ways of life across the full range of human activities, including social, educational, religious, recreational and economic life, encompassing both public and private spheres.*[25]

Kymlicka argues that societal culture is a necessary prerequisite for individuals to live a full life, and that the realization of societal culture mandates the participation of public institutions. In other words, individuals should expect public institutions to not only safeguard their interests against unjust infringement, but to also support their culturally determined way of life.

What does all this entail for public policies on minorities? Schematically, Kymlicka develops his arguments on two levels. First, he substantiates the proposition that the liberal ideal of the individual's autonomy entails the need for the individual's participation in a "societal culture." On the second level, Kymlicka asks whether it matters which "societal culture" will provide a given individual, who be-

[20] Charles Taylor, "The Politics of Recognition", in *Multiculturalism. Examining the Politics of Recognition*, edited by Amy Gutmann, 25 – 74. (Princeton N.J, Princeton University Press, 1994).
[21] See Charles Taylor 66
[22] See Charles Taylor 33
[23] Charles Taylor 25
[24] Charles Taylor 26

[25] Will Kymlicka, , *Multicultural Citizenship. A Liberal Theory of Minority Rights* (Oxford, Clarendon Press, 1995) 76

longs to a minority group, with the context which shapes that individual's choice of a certain way of life. In other words, does it matter whether minority members will model their way of life after their own national or ethnic culture, or whether they will resort to the "services" of the "mainstream" one?[26]

Kymlicka answers this question on two levels, too. First, he offers several arguments in support of the thesis that the individual's self-identity is important and valuable to her; therefore, the possible decision to assimilate into the societal culture of another ethnic group (that is, of the majority) by adopting it as her/his own culture, is not of the same order of importance as the decision to change her/his home or job.[27] Then he formulates the thesis that *justice* requires allowing the individual to freely decide whether to do one thing or another– that is, whether to assimilate or to seek ways to satisfy her/his need for a societal culture, starting from her/his existing self-identity. Or, in other words, to fight for group-specific rights for her/himself, as well as for the other members of her/his group, so that they could use those rights to collectively develop their own minority societal culture:

> *I believe that, in developing a theory of justice, we should treat access to one's culture as something people can be expected to want, whatever their more particular conception of the good. Leaving one's culture, while possible, is best seen as renouncing something to which one is reasonably entitled.*[28]

So, it is by no means impossible to formulate arguments in favor of the value of minority identities, which are in accordance with liberal values and the principles of liberal democracy in general. But what about the application of multiculturalism in reality – as a methodology in designing minority policies? The development of multicultural practices in some Western-type democratic societies has shown that policies, which support minority identities, can lead to self-segregation of minority communities[29]. Efforts to preserve a specific way of life, as well as resistance to the temptations of assimilation, can result in the distancing of the members of a minority group from mainstream society in the country they live in, prompting them to give priority to their own culture-specific values and norms over the civic values and norms governing public life in the country they inhabit. Such efforts could also lead to limiting the freedoms of the individual members in the name of keeping archaic cultural traditions alive[30]. All this has led to widespread disappointment in the potential of multiculturalism, both as theory and practice, and to a decline of its prestige and political influence.

It is not easy, of course, to explain why this has happened, why the well-intentioned (in the spirit of liberal democracy) attempts to pay due tribute to minority identities has had such a negative impact on the integrity of some of the societies where multiculturalist policies were pursued. Here I'll present just a couple of suggestions. In my opinion the main flaw of these policies was not that they were too communitarian. As I have tried to show, normative multiculturalism has very little in common with communitarianism. I see a problem rather with the *essentialist* understanding of cultural diversity, which is characteristic of the main multiculturalist conceptions. A key formulation in the multiculturalist discourse throughout the last decade of the XXth century and the first decade of the XXIst was "recognition of the value of cultures". Here "culture" is understood in holistic terms as a discrete entity. In this way the internal diversity of cultural activities is ignored. The same is true also concerning religious, ethnic, racial, and national identities. A conservative, even "conservationist" approach to them was one of the factors due to which multiculturalist policies in many cases produced self-

[26] See Will Kymlicka 84

[27] As an example, here are two of these arguments, which Kymlicka has borrowed from eminent exponents of liberal thought. Cultural identity gives individuals a sense of security because it connects them to a group of other people through relationships of solidarity based on belonging, not accomplishment (see Avishai Margalit, and Joseph Raz, "National Self-Determination." *Journal of Philosophy* 87, no 9 (1990): 439 – 461, 447). Cultural membership adds additional meaning and value to our actions, which become not only acts of individual accomplishment, but also part of a collective, historically significant, continuous creative effort whereby culture is made and remade (see Yael Tamir, *Liberal Nationalism*. (Princeton NJ: Princeton University Press, 1993) 72.)

[28] Will Kymlicka 86

[29] See Charles Guzzetta, "Should Pograms and Service Delivery Systems Be Culture-Specific in Their Design? No." In *Controversial Issues in Multiculturalism*, edited by Diane de Anda, 62 – 78. (Boston: Allyn and Bacon, 1997), 63

[30] See Wynetta Devore, "Has the Focus on Multiculturalism Resulted in Inadequate Attention to Factors such as Gender, Social Class and Sexual Orientation? Yes". In *Controversial Issues in Multiculturalism*, edited by Diane de Anda, 39 – 61. (Boston: Allyn and Bacon, 1997) 41; Also, Charles Guzzetta, 66

segregationist trends in the life of minority communities. Moreover, the encouragement of multicultural tolerance and pluralism throughout the period in question was not combined with measures to activate and elaborate communication between such groups and society at large. The importance of public communication in this regard was definitely underestimated.

So, finally: in my opinion a constructionist approach to minority identities, i.e. treating them as created and reproduced over time by deliberate activities and interactions – both by members of minority groups and various organizations, and even by individuals with ingroup membership in the majority group in society at large–complemented by two-way flows of public communication (via which claims concerning the status of these communities in society are being legitimized) might "do the job". In other wordsthis approach might help the "reconciliation" between cultural diversity and liberal democracy in more palpable and meaningful ways than a fruitless fluctuation of minority policies between universalist and particularist methodologies.

References

Deveaux M (2006) *Gender and Justice in Multicultural Liberal State* .Oxford, Oxford University Press.

Devore W (1997) Has the Focus on Multiculturalism Resulted in Inadequate Attention to Factors such as Gender, Social Class and Sexual Orientation? Yes. In *Controversial Issues in Multiculturalism*, ed. Diane de Anda, 39–61. Boston: Allyn and Bacon.

Divjak S (2018) Communitarianism, Multiculturalism and Liberalism, *Balkan Journal of Philosophy* 10(2): 147–163, 161

Guzzetta C (1997) Should Pograms and Service Delivery Systems Be Culture-Specific in Their Design? No." In *Controversial Issues in Multiculturalism*, ed. Diane de Anda, 62–78.(Boston: Allyn and Bacon.

Iskandar A, Rustom H (2006) From Paris to Cairo. Resistance of the Unacculturated. *The Ambassadors* 9(1). http://www.ambassadors.net, accessed December 2nd 2017.

James M R (2004) *Deliberative Democracy and the Plural Polity.* Lawrence, KA: University Press of Kansas.

Kanev K (1998) Law and Politics on Ethnic and Religious Minorities. In *Communities and Identities in Bulgaria*, ed. Anna Krasteva, 55–95. Ravenna, Longo Editore.

Kymlicka W (1995) *Multicultural Citizenship. A Liberal Theory of Minority Rights*(Oxford: Clarendon Press.

Laurence J (2003) The New French Minority Politics. The Brookings Institution http://www.brookings.edu/fp/cusf/analysis/laurence.pdf, accessed April 28, 2017

MacIntyre A (1988) *Whose Justice? Which Rationality?*(Notre Dame: University of Notre Dame Press.

Margalit A, Raz J (1990) National Self-Determination. *Journal of Philosophy* 87(9): 439–461, 447

Rawls J (1993) *Political Liberalism.* New York: Columbia University Press.

Rawls J (1997) The Idea of Public Reason Revisited. *The University of Chicago Law Review*, 64(3): 765–807.

Robson R (2018) In What Sense is Multiculturalism a Form of Communitarianism. *Balkan Journal of Philosophy*, 10(2): 145–146.

Sandel M J (1984) The Procedural Republic and the Unencumbered Self," *Political Theory* 12(1): 81–96

Tajfel H (1978) *The Social Psychology of Minorities*. London, Minority Rights Group.

Tamir Y (1993) *Liberal Nationalism*. Princeton NJ: Princeton University Press, 72.

Taylor C (1994) The Politics of Recognition. In *Multiculturalism. Examining the Politics of Recognition,* ed. Amy Gutmann, 25–74. Princeton N.J.: Princeton University Press.

Valadez J (2001) *Deliberative Democracy, Political Legitimacy and Self-Determination in Multicultural Societies.* Oxford: Westview Press.

Várady T (2001) On the Chances of Ethnocultural Justice in East Central Europe. In *Can Liberal Pluralism Be Exported? Western Political Theory and Ethnic Relations in Eastern Europe*, eds Will Kymlicka, Magda Opalski, 135–149. Oxford: Oxford University Press.

Waldron J (1995) Minority Cultures and the Cosmopolitan Alternative. In *The Rights of Minority Cultures*, ed. Will Kymlicka, 93–122. Oxford: Oxford University Press.

Walzer M (1994) *Thick and Thin. Moral Argument at Home and Abroad.*(Notre Dame: University of Notre Dame Press.

BALKAN JOURNAL OF PHILOSOPHY. Vol. 11, Issue 2, 2019

Book reviews

LOGICAL ATOMISM: A HUNDRED YEARS LATER

[A Review of *The Philosophy of Logical Atomism*, 2018, by Nikolay Milkov]

Denitsa Zhelyazkova

Phd Student in Contemporary philosophy
Sofia University
15, Tsar Osvoboditel Blvd.
Sofia, Bulgaria
e-mail: d_zhelyazkova@hotmail.com

The pages below discuss an impressive analytical text, namely Nikolay Milkov's book *The Philosophy of Logical Atomism*, published in 2018 by 'Science', one of the most prestigious Russian publishers of scientific literature. To those who are well versed in the biography of Professor Milkov, the publication of the book comes as no surprise. In 1983 he defended in Moscow his doctoral dissertation under the title *Critical Analysis of the Philosophy of Logical Atomism*. The dissertation is now published in Russia in book form after more than thirty-five years. Beginning in 1983 Milkov worked as a research fellow at the Institute of Philosophy to the Bulgarian Academy of Sciences. After 1989, he worked at several German universities and in 2009 was promoted to the position of full professor at the University of Paderborn (Germany), where he continues to teach.

To understand the professional biography of Nikolay Milkov is to understand the dominant place that the philosophical and logical views of Russell and Wittgenstein held for him. This is evidenced by the many books and papers he has authored, the numerous conferences and seminars he has attended, and also by his translations of philosophical texts[1]. In the first pages of the revised version of his dissertation Milkov confesses to his readers that it is only now, when preparing the text for print, more than thirty-five years later, that he has come across a considerable number of ideas in it which were developed extensively in his later works.

The book has preserved the typical structure of a dissertation: Introduction, two chapters, and a Conclusion. Each of the chapters is devoted to the two versions of logical atomism – Russell's (the first chapter) and Wittgenstein's (the second one). Russell's theory of logical atomism takes up more space in the book for one single reason: Russell's work represents the beginning of the philosophy of logical atomism. Therefore, through revealing Russell's versions of logical atomism, Milkov outlines the genesis of the idea. Russell tried out various directions of its application in search for solutions of logical, epistemological, and ontological problems that often turn out to be inseparable from each other. It is well known that Russell's logic is largely grounded and dominated by his ontological and epistemological views.

Let me try to clarify this last point. The first chapter of Milkov's book outlines the philosophical development of Russell. Following Moore's turn to realism in the end of the 19th century, Russell explicitly professes "Platonic atomism" or "Platonic realism". In the history of philosophy this name denominates the specific period of Russell's philosophical development at that moment. Milkov re-

[1] Nikolay Milkov is the first translator of Wittgenstein's Tractatus Logico-Philosophicus. See Wittgenstein, L. 1988. *Selected Works*, Trans. By Nikolay Milkov, Sofia: 'Science and Art'.

constructs the intricate picture[2] of the nature of Platonic atomism, understood as an ontological atomism which assumes that everything has a being, and everything that has a being has the status of a term[3]. The latter may be more clearly set out in one of the paragraphs of *The Principles of Mathematics* (1903):

Whatever may be an object of thought, or may occur in any true or false proposition, or can be counted as one, I call a term. This, then, is the widest word in the philosophical vocabulary. I shall use as synonymous with it the words unit, individual, and entity. The first two emphasize the fact that every term is one, while the third is derived from the fact that every term has being, i.e. is in some sense. A man, a moment, a number, a class, a relation, a chimaera, or anything else that can be mentioned, is sure to be a term; and to deny that such and such a thing is a term must always be false. (Russell, 1937: IV)

What is important for Russell is that with our propositions we should be able to speak about anything whatever. This points in the direction of the 'unbounded variable'. It becomes obvious, therefore, that logic and ontology are directly interlinked. Unlimited variables cover absolutely all entities - relationships, masses, numbers, spatial points, propositions, classes, propositional functions, etc. By considering 'x' to be always an unbounded variable we may talk about the variable that is conceptually identical in logic, arithmetic, geometry, etc. (Russell, 1937, 89; §88). Any restriction on the variable should be declared; then the truth of the statement using the restricted variable is not absolute, but rather depends on the statement restricting the variable. The variables in logic cover everything that has existence, and the absolute truth of logic and mathematics requires absolutely un-restricted variables.

Why did Russell hold these strange and extraordinary views (which he soon came to severely criticize and altogether reject)? We find the answer on page 34 where Milkov emphasizes:

> Russell believed that it is exactly the adoption of many terms which constitutes the alternative to logical monism. Therefore, this first form of atomism is the most consistent form of "vertical atomism", which asserts the independent coexistence of the "layers of being".[4] (Milkov, 2018, p.34).

In 1903, especially after writing *The Principles of Mathematics*, Russell believed he was on the way to ridding himself of all obstacles to showing that mathematics can be really true. Russell, however, also anticipated the big problems that were soon to appear. As Griffin rightly notices, since Russell does not observe the restrictions of the typological classification, he is able to produce 'left of is human' as a legitimate proposition. (Griffin, 1980, p.121). The situation is adequately summarized as follows: 'Everything … leads to the conclusion that the excessive hospitality of the "early" Russell's ontological project will accumulate liabilities, which will eventually lead at first to a repair, and then to a radical change of the Platonic atomism" (Asparuhov, 2018, p.54).

*

The core beliefs of this *first form* of logical atomism held firm in Russell's mind until 1913 when his ambition drove him to create a considerable manuscript of about 300 pages–*The Theory of Knowledge*–which was never published prior to his death. Milkov reveals the reason for this on page 71: Russell was stunned by the criticism he received from Wittgenstein after reading the manuscript. The blow had been so severe that the writing was not only abruptly halted, but the whole system of logical atomism was shaken. This episode, however, affected Russell altogether positively since it served as an impetus to go beyond his previous views and elevate them to a new level. Thus emerged the *second* logical atomism which was set forth in two of Russell's most famous books: *Our Knowledge of the External World* (1914) and *The Philosophy of Logical Atomism* (1918).

In the text *of 1913* Russell chose to ask a different question as the starting point for doing philosophy, namely, '*What differentiates a true proposition from a false one*?' (Milkov, 2018, p.71). This new turn of thought meant that, contrary to the case of the first logical atomism–with the theory of denotation as its heart–the second logical atomism took *the theory of truth* to be at its base.

This new mode of thinking placed *proposition, belief,* and *fact* as central concepts of the modified

[2] For more on the 'turn to realism' in the early analytical philosophy, see Beshkova, A. 2010. See also Lozev, 2011: especially Ch. 5, describing the 'abandonment of both Kant and Hegel' (Lozev, 2011, p.102)

[3] In Russell's philosophy 'term' is an ontological, not a linguistic category.

[4] In the recent philosophical literature in Bulgaria, the motives and the intentions of this first form of Russell's logical atomism have been analyzed in: Lozev, 2015; 2012 and Asparuhov, 2007.

version of logical atomism. In the ensuing paragraphs illustrating Russell's views of this period, Milkov explicates how *the theory* has acquired the leading role and how, as a result of this, the nature of philosophical investigation had been changed. Identification and clarification of 'logical form' now turns out to be the primary task of philosophical logic. As Milkov notices (Milkov, 2018, p.181), Russell is in agreement with Wittgenstein that philosophical logic is a 'taxonomical science' along the same line as zoology.

Milkov ends the first part of the book, dedicated to Russell's logical atomism, with a paragraph on Russell's views after 1919–that is, his 'turn to subjectivism' (Milkov) and the famous 'neutral monism' that marks a new period in Russell's philosophical development.

*

The second chapter of the book displays Milkov's analysis of Wittgenstein's version of logical atomism. It should be stressed that Milkov's reconstruction of Wittgenstein's earlier philosophy is based not only on the Tractatus, which obviously is the main source, but also on other writings, such as letters, diaries, notes, etc. This helps to elucidate the project undertaken and the route followed by the young Wittgenstein in his first philosophy.

Milkov sets the scene of his analyses by pointing out that Wittgenstein shared Frege's thesis that thought can be reached only through the analysis of language. Any abuse of the thesis will likely result in psychologism and subjectivism. Taking this thesis to be fundamental for philosophical logic quickly leads Wittgenstein to realize the importance of the relationship between logical symbol and reality. His theory of symbols gives pride of place to precisely this relationship.

Reality, according to Wittgenstein, can be *reached* and *given* to us in three basic ways: (i) by means of propositions – this is what the natural sciences do; (ii) by means of pieces of art; (iii) by means of mystical contemplation of the world as a whole – this is what religion does (Milkov, 2018, 183). Wittgenstein sees his task in philosophy to be the logical analysis of these three fundamental modes of presenting reality to the human mind.

This enables Milkov to 'structure' the investigation of Wittgenstein's logical atomism discerning two aspects of it: (a) analysis of the 'logical' *given-ness* of reality, and (b) analysis of the 'extra-logical' (non-logical) *given-ness* of reality. Thus, the next two paragraphs of the second chapter, namely 2.2 and 2.3, discuss (a) and (b) in turn.

In 2.2 Milkov discusses Wittgenstein's theories of symbols and forms. Fundamental for both of them is a theory that Wittgenstein formulated in the later part of 1913 and that Russell found to possess elements of undisputed mysticism. This theory, based on the dichotomy of 'said (expressed)' /gesagt/ and 'shown' /gezeigt/, so profoundly influenced Wittgenstein's thought until his second philosophy that it echoed throughout his thinking – even in Wittgenstein's letters to potential publishers of the Tractatus (Milkov analyzes in detail part of Wittgenstein's famous letter to Ludwig fon Ficker on p. 157). The importance of the 'said - shown' theory was that it served as a tool in Wittgenstein's hands to criticize both Russell's and Frege's logics, as well as all of traditional logic.

Milkov outlines the content of both Wittgenstein's theories – of symbols and of forms – surveying the effects in logic produced by this strange but strong theory under the spell of which Wittgenstein lived for a considerable part of his life.

Milkov dedicates the last paragraph of the second chapter (2.3) to Wittgenstein's views on ethics, aesthetics, and religion. Here the 'extra- (or rather the 'non-) logical' *given-ness* of reality is analyzed once again in reference to the 'said – shown' opposition. Here, Milkov is concerned with the most famous Wittgensteinean view of philosophy as specific *activity*.

Milkov also reveals Wittgenstein's thesis that the two modes – ethical and aesthetical – through which reality may be 'reached' are in fact 'one'. This is because they are both aspects of the *mystical*: in the first case we contemplate over *the world as a whole*, while in the second we contemplate over the objects in the world *from the perspective of eternity*.

Milkov dedicates the very last pages of the book to 'the way ahead' and elucidates the specific points of the philosophical doctrine of earlier Wittgenstein that in the 1930s and 1940s would become the sources and reasons for his theoretical crisis, and through which he would somehow 'trace' his way forward to what Russell calls 'Wittgenstein II', or the *later* (or the *second*) *philosophy* of Ludwig Wittgenstein.

Upon reading the last page and closing the book one is filled with the satisfying sense that he had really got a complete idea of what 'logical atomism' as a philosophical doctrine was, and why it used to

be so important and dominant at the end of the Nineteenth and in the first half of the Twentieth century.

References

Anscombe G E M (1971) *An Introduction to Wittgenstein's Tractatus.* 1957. Philadelphia: University of Pennsylvania Press.

Asparuhov V (2018) Russell: Philosophical Logic and the Problem of Logical Form, Sofia: Bulged.

Asparuhov V (2007) On Philosophical Logic in *What is Logic?*, 1912, *Philosophy* 3: 68–83.

Beshkova A (2010) The Turn to Realism and the Emergence of Analytical Philosophy, *Critique and Humanism (Analytical Philosophy)* 31(1): 37–47.

Griffin N (1980) Russell on the Nature of Logic (1903–1913), Synthese, Vol. 45, No. 1, Bertrand Russell's Early Philosophy, Part I (Sep., 1980), Springer, 117–188.

Lozev K (2011) *The Earlier Russell: Life and Philosophy*. Sofia: Quanta.

Lozev K (2012) Francis Bradley and His Influence on Moore and Russell. *Yearbook of the Philosophical and Political Sciences Department* 2: 54–68.

Lozev K (2015) Bertrand Russell's Logical Atomism. *Yearbook of the Philosophical and Political Sciences Department* II(1): 134–149.

Milkov N (2018) *The Philosophy of Logical Atomism.* St. Petersburg: Science.

Russell B (1937) *The Principles of Mathematics. Second edition.* London: George Allen & Unwin Ltd.

Wittgenstein L (1988) Tractatus Logico-Philosophicus, In *Selected Works*, trans. Nikolay Milkov, 49–133. Sofia: Science and Arts.

Wittgenstein L, Ogden C K (1973) *Letters to C. K. Ogden; with Comments on the English Translation of the Tractatus Logico-Philosophicus by Ludwig Wittgenstein; F. P. Ramsey; G. H. von Wright.*

Zhelyazkova D (2018) The Ontological Motivations of the Theory of Denoting: Before and after Quine. *Philosophical Alternatives* 6: 54–55.

De Regt, Henk W, *Understanding Scientific Understanding* (NY: Oxford University Press, 2017, 320 pp.)

Rossen Naydenov Stoyanov
PhD candidate at SU "St. Kliment Ohridski", Sofia, Bulgaria
e-mail: rstoyanov@phls.uni-sofia.bg

1. Introduction

It is only recently that "understanding" has become a topic of debate within the field of the philosophy of science. The reason for this "delay" has been the philosophers' focus on "scientific explanation" as the proper notion for analysis, a preference that can be traced back to Carl Hempel's and Paul Oppenheim's 1948 "Studies in the Logic of Scientific Explanation" (repr. in Hempel, 1965). This pioneering paper not only marked the beginning of extensive discussion of explanation in science, but it also dismissed "understanding" as an overly psychologistic and pragmatic notion, and thus steered philosophical attention away from it. Admittedly, "scientific understanding" has not been neglected completely – first it was Michael Scriven's "Explanations, Predictions and Laws" (1962) that brought the notion to the fore. Later, in his 1974 "Explanation and Scientific Understanding" Michael Friedman argued that philosophical accounts of explanation should clarify how such explanation produces understanding. Since then philosophers have often pointed out the understanding-providing power of their preferred models of explanation. However, apart from these few exceptions, as well as Achinstein's *The Nature of Explanation* (1983), philosophical attention has remained focused on "explanation" and detailed analysis of "understanding" has been avoided.

It was in the field of epistemology with the work of Linda Zagzebski (1996), Catherine Elgin (1996), and Jonathan Kvanvig (2003) that a discussion was provoked about the difference between epistemology's traditional subject matter of knowledge and understanding; as a result, the position of the latter was established as a cognitive achievement worthy of philosophical examination. As far as the philosophy of science is concerned Henk de Regt's work has been among the main factors for reintroducing "understanding" and its relations with "explanation" as one of the issues of discussion within the field. Thus, in the last ten years there has been an increase in related research, with an attendant increase in the number of publications on this topic : *Scientific Understanding: Philosophical Perspectives* (de Regt, Leonelli and Eigner, 2009), *The Nature of Scientific Thinking: On Interpretation, Explanation, and Understanding* (Faye, 2014), *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science* (Grimm, Baumberger and Ammon, 2017), *Understanding, Explanation, and Scientific Knowledge* (Khalifa, 2017) as well as de Regt's own *Understanding Scientific Understanding* (2017).

2. An HPS approach

The composition of *Understanding Scientific Understanding* – it is divided into two major parts, a theoretical part (chapters 2 to 4) and a historical one (chapters 5 to 7) – reflects its author's ambition: unlike mainstream philosophical discussions of explanatory understanding, which de Regt qualifies as historically uninformed, his inquiry aims to combine successfully systematic philosophical analysis with detailed historical narrative. He explicitly positions his work in the tradition of integrated history and philosophy of science approaches (HPS). However, as de Regt clarifies in the introduction, this structure does not reflect a relation between theory and historical record, in which the latter provides an empirical test for the former. In contrast to this "confrontational model of HPS," de Regt claims to adopt an approach, which, following suggestions by Jutta Schickore (2011) and by Hasok Chang (2012), proceeds upon a close and constant interaction between philosophical concept analysis and historical narrative whereby philosophically clarified notions are used as tools for approaching and interpreting historical data, but also historical data help to further refine and modify the initially articulated notions.

Thus, as de Regt points out, it was a "historiographical puzzle", namely, the controversy at the genesis of quantum mechanics in the 1920s, that inspired his theory of scientific understanding since the mainstream philosophy of science, as he saw it, appeared incapable of providing an adequate account of the research strategies adopted by the opposing parties involved. This case study, as well as another one included in the historical part of *Understanding Scientific Understanding*–the controversy between Maxwell and Boltzmann over the specific heat anomaly–can be traced back to de Regt's Ph.D. thesis "Philosophy and the art of scientific discovery" (1994). Focusing on these two controversies, he thereexplicated the disagreement, as well as the particular solutions scientists espoused, in reference to their philosophical positions. This interpretive move implied the rejection of the traditional distinction between the context of discovery and the context of justification. *Understanding Scientific Understanding* represents the integration of this historical work with de Regt's most recent elaboration of his contextual theory of scientific understanding. This theory was first presented in "Reduction and Understanding" (de Regt and Dieks, 1998) and "Ludwig Boltzmann's Bildtheorie and Scientific Understanding" (de Regt, 1999), then it was more fully developed in "A Contextual Approach to Scientific Understanding" (de Regt and Dieks, 2005), and further refined in "Understanding and Scientific Explanation" (in de Regt, Leonelli and Eigner, 2009).

Completely in the line of the HPS approach, what *Understanding Scientific Understanding* aims at is a theoretical account of scientific understanding that pays due attention to scientists' beliefs and to actual scientific practice. That is why, de Regt argues, it is a contextual theory – one that accounts successfully for historical and disciplinary variation in the standards of explanatory understanding, and yet still specifies some general criteria for achieving scientific understanding whereby the purely local context is transcended. History has shown, the author insists, that the search for universally valid, timeless standards is futile.

3. Understanding Scientific Understanding – the theory

De Regt builds the book's theoretical component by taking three steps. First, he clarifies and rejects the reasons behind what he sees as a neglect of the notion of understanding by philosophers of science (chapter 2). Then he examines accounts of understanding provided by standard models of scientific explanation and argues against the universal validity of any one of them; what is needed, he claims, is an overarching pluralist theory that explains how each of these can provide understanding (chapter 3). Finally, he formulates his own contextual theory of scientific understanding (chapter 4).

3.1 Against the subjective nature of understanding

As mentioned, Oppenheim's and Hempel's work was of huge significance for the subsequent exclusive focus of philosophers of science on the notion of explanation, and thus, it is not simply by chance that de Regt confronts the main charge usually leveled against any attempt for a philosophical account of "scientific understanding", that of pragmatism and subjectivity, as it was articulated by Hempel (1965, 1983).

In accordance with the spirit of logical empiricism which, according to de Regt, remained dominant in the field of scientific explanation in the 1960s and the 1970s, Hempel took "scientific understanding" and "intelligibility" as belonging not to the field of philosophy, but to psychology or the history of science – as scientific laws and theories are objective in the sense that their empirical implications and evidential support are independent of scientists' beliefs and attitudes, the aim of the philosophy of science should also be the objective analysis of scientific statements and their relation to evidence. In contrast, by means of their reference to human subjects, the explainers and the audience receiving the explanation, the notions of intelligibility and understanding are pragmatic and, hence, relative, since something can be intelligible for one audience and unintelligible for another. This equation of the pragmatic with the subjective and the psychological is shared by J. D. Trout (2002, 2007) who draws attention to the psychological biases that may affect the so-called aha-feeling of understanding and that is why, according to him, the feeling of understanding is unreliable when the correctness of an explanation is being evaluated.

De Regt agrees that "understanding" can be used only in reference to human agents – "understanding" is a three-term relation between a phenomenon, an explanation, and the person who uses the explanation in order to attain understanding of the phenomenon. He also distinguishes between the feeling of understanding (FU), object of Trout's criticism, the understanding of a phenomenon (UP), which is usually regarded as the main epistemic aim of science, and the understanding of a theory (UT)

equated by him with an agent's ability to use the theory. De Regt is cautious to avoid confusion between UP and UT and he renders "S understands theory T" as "T is intelligible to S", with the "intelligibility" of a theory being an important notion in his theory.

De Regt interprets the position shared by Hempel and Trout as based on the adoption of a strict division between the epistemic and the pragmatic, where the former refers to the beliefs of scientists about the relation between theory and world, whereas the latter refers to their attitude towards the theory itself. This strict division, claims de Regt, presupposes that the epistemic status of a theory depends on a direct evidential relation with phenomena.

Thus, in his refutation of Hempel's influential claim asserting the irrelevance of "understanding" to the philosophy of science, de Regt follows two steps. First, he argues against the above mentioned division by showing that the evidential relation between theory and phenomena is not direct, that is, there is no "objectivity" as independence from scientists' beliefs and attitudes. Second, following Friedman's suggestion that "objective" can be understood as "being constant for a large class of people" he refers to Longino's and Douglas's analyses of the notion of "objectivity" which demonstrate that its being dependent on the beliefs and attitudes of scientists does not make it "subjective" in the sense of varying from one individual to another (Longino, 1990; Douglas, 2009a, 2009b).

The lack of direct evidential relation is clear, according to de Regt, in scientific practice, where skills and scientific models mediate between scientists and phenomena. Skills and models are necessary for the construction of scientific explanations, which de Regt takes to be arguments in the broad sense – answers to the question of why a particular phenomenon occurs, which present a "systematic line of reasoning that connects [them] with other accepted items of knowledge (e.g. theories, background knowledge)"; explanations are not necessarily linguistic, they may take the form of pictorial representations, as well.

Providing as examples Hempel's own deductive model, Wittgenstein's qualification of what it means to understand a mathematical series (Wittgenstein, 1953), as well as Harold Brown's discussion on constructing and evaluating deductive proofs in logic (Brown, 2000), de Regt concludes that these point to the indispensability of skills and tacit knowledge for the construction and evaluation of arguments/explanations. Moreover, these are skills that are learned through practice alone, and are, therefore, contextual. The lack of any direct evidential relation between theory and phenomena is also testified to, according to de Regt, by scientific models. He refers to Nancy Cartwright's "simulacrum account" of explanation (Cartwright, 1983) and to Morgan's and Morrison's notion of the "autonomy" of scientific models (Morgan and Morrison, 1999) and argues that models are derived neither from theory nor from empirical data. Their construction does not follow any algorithms or formal principles, but requires idealizations and approximations which ultimately rely on practical judgment and skills. Thus, de Regt claims, the achievement of a scientific explanation has an inherent pragmatic aspect.

For the clarification of the nature of these skills de Regt uses the notion of theoretical virtue and refers to Thomas Kuhn's list of five theoretical desiderata (Kuhn, 1977), that is, to the qualities of a theory which influence one's choice of theory as values rather than determine it as rules: accuracy, consistency, scope, simplicity and fruitfulness. Skills and theoretical virtues, argues de Regt, are correlated since scientists prefer theories which correspond to the skills they have acquired and improved in training and research; that is why this preference varies across disciplines and historical periods. Other theoretical virtues mentioned by de Regt are visualizability, causality, continuity, and locality, but he distinguishes as the two basic ones empirical adequacy and internal consistency, claiming that in choosing a theory there is a trade-off between these basic virtues and the others. The value that scientists attribute to the cluster of qualities of a theory that facilitates its use de Regt defines as "intelligibility" of a theory.

This recognition of the significance of theoretical virtues in theory choice is reflected in the way the notion of "objectivity" is reformulated by Longino and Douglas, who argued that a "strong", value-free objectivity of the kind adopted by Hempel, is unattainable and unrealistic. As far as "objectivity" refers to the individual scientist's thought processes de Regt adopts Douglas's notion of "detached objectivity," according to which evidence should not be disregarded in favor of theoretical virtues, but rather these virtues should be consulted when evidence does not speak unambiguously. On the other hand, in social processes objectivity is established, as Longino has argued, either when there is agreement within the community (concordant objectivity) or when–after an initial disagreement–consensus is

achieved by means of critical discussion (interactive objectivity).

This reformulated notion of "objectivity" allows de Regt to claim that his account of understanding is not "subjective" in the sense of "varying across individuals," which was used by Hempel to dismiss the philosophical relevance of "understanding," but is, in fact, "objective" in the sense of being "detached", "concordant" and "interactive," and hence "varying across scientific communities and historical periods". Such variability is something, which, as de Regt insists, corresponds to the reality of scientific practice and to the testimony of historical record.

3.2 The variety of the models of explanation

Variability in the ways understanding is delivered is characteristic also of the models of explanation established within the field of the philosophy of science. De Regt's critical examination of these (chapter 3) aims to show that each of them relies on plausible intuitions of how understanding is achieved, but none of them can claim the position of a universal account. Thus, though unificationism (Hempel, Friedman, Kitcher) does not pose restrictions on the form of understanding-providing theories, the historical record shows that causality and visualizability have also been considered as essential for achieving understanding. Wesley Salmon's attempt to establish causation as the privileged road to scientific understanding also faces difficulties, according to de Regt – he points to domains of science where causal analysis is not applicable (quantum theory) or where though prevalent, it may not always be preferred (biology); finally, sometimes it is applicable in principle but yet not preferred by scientists (the explanation of the special relativistic Lorentz contraction). De Regt refers to the alternative causal models of explanations such as Woodward's and the ones of the new mechanists (Machamer, Darden, Craver). De Regt draws attention to the diversity in all these commonly labeled "causal" approaches – whereas Salmon and Strevens claim that causal analysis reveals hidden structures of the world (in Salmon, 1998; Strevens 2008), Woodward concludes that understanding is achieved not when one gets the ontology right, but when one knows the behavior of systems (Woodward, 2003); and, finally, the new mechanists' approach employs not only causality, but also visualization and locality (Machamer, Darden, and Craver, 2000).

De Regt also examines Salmon's attempt to reconcile the causal-mechanical and the unificationist models (Salmon, 1998), as well as Railton's "ideal explanatory text" thesis (Railton, 1981). Salmon's attempt, de Regt argues, does not solve the problems with quantum theory and with the Lorentz contraction, while Railton's thesis implies that there are objective, universal, and immutable criteria of genuine understanding, a notion that conflicts with the historical record and with scientists' own beliefs and attitudes – understanding, then, may be achieved by various types of explanations.

De Regt's conclusion is that though all these various accounts reflect our intuitions of what understanding consists of, an overarching theory is required that can accommodate them as different legitimate roads to the achievement of understanding. He claims that his theory successfully meets this requirement.

3.3 De Regt's contextual theory – CUP and CIT

Having defined 'intelligibility' as the value that scientists attribute to the cluster of qualities of a theory that facilitates its use, de Regt formulates two other basic notions: criterion for understanding a phenomenon (CUP) and criterion for the intelligibility of a theory (CIT). Thus, according to him, a phenomenon P is understood scientifically (UP) if and only if there is an explanation of P that is based on an intelligible theory T (UT) and that conforms to the basic epistemic values of empirical adequacy and internal consistency. To possible charges that this is a "theory-dominant" approach de Regt responds that he shares Giere's more liberal view of theories as collections of principles that provide the basis for the construction of models (Giere, 2006). CUP is pragmatic since it involves the notion of intelligibility, which is pragmatic itself – whether a theory is regarded as intelligible does not depend only on the theory, but on the scientists involved and on their skills and background knowledge. Pressed to balance between the requirement for context-dependency and not to qualify "intelligibility" as a matter of taste, de Regt formulates the criterion for the intelligibility of a theory (CIT), which states that a scientific theory T is intelligible for scientists (in a particular context) if they can recognize qualitatively characteristic consequences of T without performing exact calculations. De Regt is quick to point out that the criterion concerns theories formulated in mathematical terms, as it is usually the case in the physical sciences; he clarifies that this does not imply the replacement of exact calculation because mathematical techniques and logical rigor are indeed essential in modern science.

Relating "intelligibility" to the scientists' ability to recognize qualitative consequences makes this notion a measure of how fruitful a theory is for the construction of models by scientists. The contextual nature of "intelligibility" is clear since this recognition of a theory's qualitative consequences requires an appropriate "match" between the skills of scientists and the theory's qualities (e.g. visualizability, causality, unification, etc.). Referring to the various ways understanding is provided by the standard models of explanation, de Regt characterizes the qualities of a theory as "tools" for generating such qualitative "predictions". Thus, as an illustration of the function of visualization, de Regt refers to the use of the MIT bag model, which helps hadron physicists to understand quark confinement by representing hadrons as "elastic bags". But, as de Regt admits, familiarity with and intuition for the behavior of the solutions of the mathematical equations used by scientists can also allow for qualitative predictions. Causal reasoning, on the other hand, provides understanding to the extent that it allows scientists to predict how the systems under study will behave under particular conditions and not because it reveals the underlying mechanisms of the world.

4. Understanding Scientific Understanding – the history

The book's second part consists of detailed descriptions of three episodes in the history of science that involved explicit debates on the issues of understanding and intelligibility. These are: the 17th-century debates on the acceptability of Newton's gravitational forces, the widespread use of mechanical models in the second half of the 19th-century and the controversy between Maxwell and Boltzmann over the specific heat anomaly, and, finally, the issue of visualizability in quantum physics as it was first disputed by Schroedinger and the representatives of the Copenhagen school and later, as it was utilized by Richard Feynman in his diagrammatic approach.

De Regt's first case is presented as a solution to the confusion which, according to him, characterizes the standard historiographical treatments of this same episode. On the one hand, there is the positivist approach of Ernst Mach and Philipp Frank ,who regarded intelligibility as a psychological illusion, resulting from scientists' familiarity with a theory (in Mach, 1872, 1911; Frank, 1957) – since for Mach and Frank it is only a theory's empirical success that matters, they viewed metaphysically motivated opposition to Newtonian gravity as a hindrance to scientific progress. On the other hand, James Cushing agreed that Newtonian formalism's empirical success was the reason for the acceptance of acting-at-a-distance forces, but he argued that Newtonians did not truly understand gravity; this was only achieved 200 years later by Einstein and his theory's local and causal account (Cushing, 1994).

In response to this historiographical 'standoff' de Regt applies the theoretical tools of his contextual theory and distinguishes between metaphysical intelligibility and scientific intelligibility. Thus, contrary to Mach and Frank, he points out that metaphysical intelligibility was important not only for Newton's opponents, but for Newton himself – he also tried to determine the metaphysical cause of gravity, but unlike his opponents he did not remain committed to Cartesian corpuscularism and was willing to look for alternative doctrines. However, Newton also separated methodologically metaphysics, physics, and mathematics, allowing him to develop a theory of gravitational forces that did not account for their causes – a natural phenomenon was considered explained when a mathematical model was provided. Again, contra the positivist interpretation which saw metaphysics as an obstacle to scientific progress, de Regt insists on the positive role of Cartesian metaphysics for the development of physical theories. Cartesian corpuscularism not only provided an ontology that was considered inherently intelligible, but it allowed for qualitative reasoning about phenomena in terms of visualizable mechanisms (e.g. the billiard balls collision analogy). When this was combined with a Galilean purely kinematical conception of motion the result was the mathematical (geometrical) representation of phenomena in the mechanics of Christiaan Huygens. That is how, according to de Regt, the positivist approach ignored intelligibility's epistemic significance.

On the other hand, contra Cushing's criteria of intelligibility, which he qualifies as too static, de Regt argues that Newton's theory was indeed intelligible to the ones who adopted it; allowing for qualitative reasoning in terms of forces acting at a distance, it facilitated the construction of explanations, e.g. of the tides, of the moon's motion, etc. Newtonian theory's final triumph was due to its fertility – scientists did not just grow accustomed to it, but they acquired the necessary skills in order to use it effectively in constructing explanatory models of phenomena.

One focus of the book's second case study is the instrumental use of mechanical models as tools for understanding. For famous 19th-century scientists

such as Thomson, Maxwell, and Boltzmann the realism of models was of little concern (Maxwell's position changed over time), what mattered was that models facilitated research. As concrete, tangible constructions these gave Thomson the ability to visualize and manipulate the object under investigation, while Maxwell's adoption of the approach of physical analogies led him to significant discoveries in electromagnetism; finally Boltzmann's Bildtheorie provided the framework of his pragmatic notion of what it means to understand a mechanism, namely, to be able to handle it correctly and know what effects certain manipulations would bring. Another focus of this chapter is the non-deductive character of model construction, as well as the significant role of judgment in this process, as illustrated by the controversy between Maxwell and Boltzmann over the specific heat anomaly. Neither entailed by the kinetic theory nor derived from experimental data, Boltzmann's 'dumbbell' model of diatomic molecules was the product of approximation and idealization. The significance of judgment was brought clearly forth by Maxwell's rejection of the model - his criticism focused on the issues of empirical adequacy and internal consistency, and the disagreement between the two scientists shows, according to de Regt, that they weighted these theoretical values of the model differently, with Boltzmann espousing practical success as the only criterion of correctness.

De Regt's claims about the significance of the virtues of a theory not only for its acceptance by the scientific community, but also for its successful application are illustrated by the book's last case study, which examines the issue of Anschaulichkeit (standing for both "visualizability" and "intelligibility") in the field of quantum mechanics. Schroedinger's emphasis on a theory's visualizability, that is, on its formulation within a space-time framework and, hence, his employment of wave mechanics, facilitated the adoption of his approach by the majority of physicists; they found Heisenberg's matrix mechanics much less accessible. But this also affected the theory's application – as de Regt argues, Schroedinger's approach initially proved to be more successful than Heisenberg's alternative. However, as de Regt's pluralist position implies, understanding can be achieved by following different roads. This is what is argued for by the description of Pauli's insistence on the inapplicability of the concepts of classical physics in the atomic domain, and hence, his rejection of Anschaulichkeit as visualizability via mechanical models. Pauli's conviction that new concepts, even if not visualizable, could become anschaulich as a result of scientists' becoming familiar with them inspired Heisenberg's reformulation of Anschaulichkeit as the "ability to think qualitatively through the experimental consequences of a theory [without contradictions]". This reformulation, as de Regt points out, provided the foundation for his own articulation of CIT. The function of visualizability as a conceptual tool for the construction of explanatory models is again illustrated by the episode of the discovery of electron spin – completely in the spirit of the old quantum theoretical approach, the discovery was the result of Goldsmit's and Uhlenbeck's skills for visualization. Pauli's reluctant surrender to this solution shows that he could not deny that the spin hypothesis successfully accounted for the empirical data; soon he found a way to interpret it within the framework of his own highly abstract approach. And yet, even today as de Regt argues, the concrete picture of rotation facilitates the understanding of electron spin. De Regt's account of the dispersion of Feynman's diagrams, which provided visualizable alternative to Schwinger's and Tomonaga's mathematically rigorous formulation of quantum electrodynamics, again emphasizes the significance of visualization as a valuable aid in understanding in physics. Feynman's diagrams were not realistic representations of physical processes, yet they proved to be useful tools for making calculations and solving problems–tools that required the acquisition of specific skills for their use and that proved to be applicable in contexts outside that of their original development.

5. Understanding Scientific Understanding – scope, normativity, and relativistic implications

The book's concluding part addresses three issues: that of the theory's scope, of its relativistic implications, and of its normative status. Since it is the second issue that has drawn much attention and critical fire in responses to de Regt's formulation of the theory in his earlier articles (see Bangu, 2017; Khalifa, 2012; Psillos, 2017; van Camp, 2014), here I discuss it last by drawing on de Regt's position as it is formulated also within the book's main body.

De Regt admits that he developed his theory with a focus on the natural sciences, and thus its criterion of intelligibility of theories (CIT) is designed for mathematized theories; he suggests that this may make it applicable to theories in fields such as economics and some branches of psychology and cognitive science to the extent that these are mathematically treatable. De Regt does not overlook the verstehen-erklaeren debate in the beginning of the

20th century or neglect to clarify that his theory does not argue for the privileged status of "understanding" as an approach specific for particular sciences; his approach focuses on understanding as related to explanation in general. Yet, he argues, his theory has drawn attention to the significance of understanding even for the natural sciences by highlighting the role of value and judgment in the construction and evaluation of scientific explanations. He admits that the contextual way in which understanding is achieved is reminiscent of the schema of the hermeneutic circle: first, having particular training and background theoretical commitments, scientists prefer theories with virtues that correspond to their skills and use them in order to construct explanations (UT); then, if successful (success here is due not only to theories' intelligibility, but also to their conformity to the basic scientific values of empirical adequacy and internal consistency), these explanations are taken as confirming scientists' understanding of phenomena; finally by using the same skills that were needed for the initial theoretical construction, scientists employ the explanations in further application, extension, and refinement of their knowledge, and thus improve their understanding of the field under examination (UP). The circularity involved here, claims de Regt, is inherent to scientific practice itself.

As far as the normative status of his theory is concerned, de Regt distinguishes between two types of normativity–prescriptive and evaluative–and argues that it is to the latter that his approach is committed – it does not specify rules that scientists should follow in order to achieve understanding, but it does clarify why in a given context certain theories are more successful by suggesting a test for their intelligibility: the CIT. Qualifying his approach as belonging to the HPS tradition, de Regt claims that his theory goes beyond the mere historical description – it identifies and explicates within scientific practices norms and criteria for understanding, thus constructing a general philosophical framework for the interpretation and explanation of historical episodes.

In opposition to anti-realism, which claimed that theories are merely tools for description and prediction of phenomena, realists, as de Regt points out, postulated that genuine understanding can be produced only by theories that are (at least approximately) true descriptions of an underlying reality. However, as the author of *Understanding Scientific Understanding* argues, scientific practice, as well as the history of science, with its pessimistic induction-reading, suggest that this connection between understanding and truth should be severed. As far as scientific practice is concerned this is what the widespread use of unrealistic or false models and assumptions shows – de Regt lists examples from economics, ecology, biology, and neuroscience. Thus, the aim of scientists is the manipulation of models, which allows for inferences to be made regarding the behavior of the system under investigation. This manipulation is always related to particular explanatory purposes and levels of description, and thus what matters is not that the representation of the target system is literally true, but that it suits the intended purpose. Moreover, even realistic depiction of a complex system may sometimes prove to be an obstacle to scientists' goal of predicting the system's behavior, and thus, to be an obstacle to scientific understanding. De Regt explicitly adopts Ronald Giere's perspectivist approach to scientific theories and models, according to which these can be compared to maps – like the latter, the former do not aim at providing a fully realistic picture and their truth is not unqualified objective truth, but rather truth relative to a particular context and purpose. That is why, de Regt claims, theories that we consider "outdated" or models that are incompatible can still be used by scientists, as this is the case, e.g. with the outdated Newtonian mechanics and with the incompatible models of water that represent it either as a collection of molecules (solvent) or as a continuum (fluid). Here, de Regt makes the provocative claim that in principle it is not impossible to use even rejected theories, such as the phlogiston theory, in order to achieve understanding; what makes it impossible is not that this postulated entity does not refer to anything in nature, but that there is no context in which a phlogiston explanation is empirically adequate and more intelligible than an oxygen explanation.

In response to charges in anti-realism de Regt argues that the context-dependency of understanding postulated by his theory does not entail relativism about science as a whole, since he does provide overarching criteria that indicate when understanding has been achieved. This is the criterion of understanding a phenomenon (CUP), which takes into account not only the intelligibility of a theory, but also its conformity to the basic epistemic values of empirical adequacy and internal consistency. Adopting Thomas Kuhn's response when his own approach was criticized as relativistic (in Kuhn, 1977), de Regt agrees that intelligibility varies from one scientific community to another, since the acquisition of particular skills is involved, but that it is such

values as the above mentioned, that transcend communities' paradigms, make inter-paradigmatic communication possible, and provide protection against "unbridled relativism".

It is again these values that are used by de Regt in order to deflect the criticisms launched at another pragmatic approach, that of van Fraassen – contra van Fraassen's approach, Salmon and Kitcher had argued that further specification of explanatory relevance is needed since otherwise any relation can be explanatory in a given context (Kitcher and Salmon, 1987), and de Regt's inclusion of the basic scientific values of empirical adequacy and internal consistency allegedly meets the requirement of this specification. The vulnerability of van Fraassen's approach is related, de Regt argues, to the distinction between the epistemic and pragmatic dimensions of science and to the assumption that the epistemic aspects of a situation being explained are pre-given and independent of the pragmatic aspects, that is, the concerns and interests of scientists. Thus, scientific relevance (epistemic dimension) and explanatory relevance (pragmatic dimension) seem to be separated, and van Fraassen does not specify objective criteria for the latter; hence, Kitcher's and Salmon's critique.

De Regt claims that his aim is not to defend anti-realism, but to present a more sophisticated picture of the relation between realistic representation and explanatory understanding.

6. Conclusion

De Regt's *Understanding Scientific Understanding* does not present the reader with a new, previously unknown theory, but collects its author's work on the notion of "scientific understanding" as it was developed over the years. However, the book is not a simple collection, but a presentation of de Regt's approach as a structured whole, which unfolds from a critical identification of deficiencies in the traditional dismissive qualification of "understanding" as subjective and psychologistic to a positive philosophical account of this notion.

Thus, the book allows for a clear overview of the proposed theory's main constitutive claims and of their positioning in the overall argumentative structure, which helps for the localization of points of tension, loci for a critical discussion.

One such point is the way "understanding" is "operationally defined", as van Camp qualifies it, as an ability (van Camp, 2014) – by referring to the performance of a deduction (Hempel's own deductive model), to the discovery and application of the formula of a mathematical series (Wittgenstein) and to the construction of a deductive proof in logic (Brown) – de Regt relies on our intuitions about the connection between understanding something and being able to use it. However, intuitions can pull us in different–even opposite–directions, as this is illustrated by the debate in epistemology about the relation between understanding and knowledge (particularly telling is the divergence of interpretations of Kvanvig's "Comanches' domination of the southern plains" thought experiment, in Kvanvig, 2003; Grimm, 2006, 2014; Pritchard, 2009; Greco, 2014).

Another point of tension is the notion of "objectivity," which characterizes the relation between the inquirer and the object of inquiry. This relation is presented by de Regt as mediated by the fictive instruments of models/theories and its constructed nature is at the basis of his claim of the contextuality of understanding and, hence, of its unavoidably pragmatic character. This relation is articulated as consisting of understanding a theory (UT) and understanding a phenomenon (UP). A certain ambiguity can be detected here, which is exploited by van Camp in his critique of de Regt's position. On the one hand, de Regt insists on the necessity of UT for UP and it may be argued that this necessary connection makes the borders between the two not easily distinguishable. This closeness between understanding a theory and understanding a phenomenon is evident in de Regt's comments on Humphreys distinction between primary and secondary understanding (Humphreys 2000), the former being the understanding of a phenomenon achieved by a scientist by means of research, and the latter is the understanding gained by fellow scientists when this research's results, that is, the primary understanding, is being communicated to them – according to de Regt the act of understanding a ready-made theory involves following the steps of its inventor and reproduction of the original line of reasoning. On the other hand, de Regt explicitly distinguishes UT from UP and he defines "intelligibility" operationally (as the ability of drawing qualitative consequences), whereas "explanation" is defined in a unificationist way (as argument that connects answers to a why-question with other accepted items of knowledge); thus, an explanation explains when it meets the additional criteria of empirical adequacy and internal consistency. Van Camp takes this difference as showing that in de Regt's account understanding is decoupled from explanation.

It may be argued that the distinction between understanding a phenomenon and understanding a

theory is also at the basis of the critique of Stathis Psillos, who insists that it is only the latter that is audience-oriented. Thus, the contextual nature of "understanding," as well as de Regt's claim that the link between understanding and truth should be severed, provoked qualifications of his position as "neo-instrumentalist" (Psillos) and as entailing "almost extreme subjectivism and relativism" (Bangu). The position presented in *Understanding Scientific Understanding* relies, similarly to other pragmatic approaches to "understanding", such as that of Jean Faye (2014), on the use of models in scientific practice, as well as on the pessimistic-induction reading of the history of science. That is why scientific realists who oppose non-factive accounts of understanding, such as Psillos, insist on the truth quam proxime of models and point to the continuity in the history of science, e.g. as it is illustrated in Einstein's recovery of the Newtonian theory of gravity as a limiting case of his own General theory of relativity (in Psillos, 2017). Thus, despite his claims to the contrary, de Regt's position is taken to imply anti-realism.

However, the pessimistic-induction reading of the history of science is not just an argument to which de Regt simply refers in support of his position; as the case studies in *Understanding Scientific Understanding* show, he examines closely the historical record, which both inspires his theoretical account of "understanding" (the main function of the second and the third case studies) and provides testing grounds for this notion's application (the first case study). This hermeneutical dynamic corresponds to de Regt's characterization of the process of scientific inquiry as one following the curves of a hermeneutical circle, but it may also raise questions characteristic of the debate on the HPS approach (see Pitt, 2001 and Schickore, 2011). Thus, it can be an object of historically informed debate to what extent de Regt's introduction of two types of intelligibility–metaphysical and scientific–in his study on the acceptance of Newtonian forces is not anachronistic, a result of the "intrusion" of his theoretical account of "understanding" upon his historical hermeneutic. And yet, the resulting narrative in *Understanding Scientific Understanding* seems to be characterized by greater historical sensitivity than the alternative ones of Frank, Mach, and Cushing.

Thus, *Understanding Scientific Understanding* represents a complete account of the position de Regt has been elaborating over the years, an account that draws on fields such as scientific representation, objectivity, models of explanation, scientific practice, explicit and implicit knowledge; moreover, it convincingly combines philosophical analysis with historical narrative. That is why, in my opinion, it may become a reference source for one of the pragmatic approaches in the recently reinvigorated debate on scientific understanding, but also a valuable contribution to the growing literature of HPS.

References

Achinstein P (1983) *The Nature of Explanation*, NY: Oxford University Press.

Bangu S (2017) Is Understanding Factive? Unificationism and the History of Science. *Balkan Journal of Philosophy* 9(1): 35–44.

Brown H (2000) Judgment, Role in Science. In *A Companion to the Philosophy of Science*, ed Newton-Smith N H, 194–202, Oxford: Blackwell.

Cartwright N (1983) *How the Laws of Physics Lie*, Oxford: Clarendon Press.

Chang H (2012). Beyond Case Studies: History as Philosophy. In *Integrating History and Philosophy of Science*, eds S. Mauskopf, T. Schmalts, 109–124, Dordrecht: Springer.

Cushing J (1994) *Quantum Mechanics: Historical Contingency and the Copenhagen Hegemony*, Chicago: University of Chicago Press.

De Regt H W (1993) Philosophy and the Art of Scientific Discovery, Ph.D. thesis, Vrije Universiteit Amsterdam.

De Regt H W, Dieks D (1998). Reduction and Understanding, *Foundations of Science* 3(1): 45–59.

De Regt H W (1999) Ludwig Boltzmann's Buildtheorie and Scientific Understanding, *Synthese* 119(1–2): 113–134.

De Regt H W, Dieks D (2005) A Contextual Approach to Scientific Understanding, *Synthese* 144(1): 137–170.

De Regt H W, Leonelli S, Eigner K (eds) (2009). *Scientific Understanding: Philosophical Perspectives*, Pittsburgh: University of Pittsburgh Press.

Douglas H (2009a). Reintroducing Prediction to Explanation, *Philosophy of Science* 76(4): 444–463.

Douglas H (2009b). *Science Policy, and the Value-Free Ideal.* Pittsburgh: University of Pittsburgh Press.

Elgin C (1996) *Considered Judgment.* N.J: Princeton University Press.

Faye J (2014) *The Nature of Scientific Thinking: On Interpretation, Explanation, and Understanding.* London: Palgrave MacMillan.

Frank P (1957) *Philosophy of Science: The Link Between Science and Philosophy*, Englewood Cliffs, NJ: Prentice-Hall.

Friedman M (1974) Explanation and Scientific Understanding, *Journal of Philosophy* 71(1):5–19.

Giere R (2006) *Scientific Perspectivism*. Chicago: University of Chicago Press.

Greco J (2014) Episteme: Knowledge and Understanding, in *Virtues and Their Vices*, eds K.Timpe, C. Boyd, 285–301, Oxford: Oxford University Press.

Grimm S (2006) Is Understanding a Species of Knowledge? *British Journal for the Philosophy of Science* 57(3): 515–535.

Grimm S (2014) Understanding as Knowledge of Causes. In *Virtue Epistemology Naturalized: Bridges between Virtue Epistemology and Philosophy of Science*, ed. A. Fairweather, 329–345, Dordrecht: Springer.

Grimm S, Baumberger C, Ammon S (eds) (2017) *Explaining Understanding: New Perspectives from Epistemology and Philosophy of Science*. NY: Routledge.

Hempel C G (1965) *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* NY: Free Press.

Hempel C G (1983) Valuation and Objectivity in Science. In *The Philosophy of Carl G. Hempel*, ed. J Fetzer, 372–396, NY: Oxford University Press.

Humphreys P (2000) Analytic versus Synthetic Understanding. In *Science, Explanation, and Rationality: aspects of the philosophy of Carl G. Hempel*, ed. J. Fetzer, 267–286, Oxford: Oxford University Press.

Khalifa K (2012) Inaugurating Understanding or Repackaging Explanation? *Philosophy of Science* 79(1): 15–37.

Khalifa K (2017) *Understanding, Explanation, and Scientific Knowledge.* Cambridge University Press, UK.

Kitcher P, Salmon W (1987). Van Fraassen on Explanation, *Journal of Philosophy* 84(6): 315–330.

Kuhn T (1977) Objectivity, Value Judgment and Theory Choice. In *The Essential Tension*, 320–339, Chicago: University of Chicago Press.

Kvanvig J (2003) *The Value of Knowledge and the Pursuit of Understanding*, NY: Cambridge University Press.

Longino H (1990) *Science as Social Knowledge: Values and Objectivity in Scientific Inquiry*, NJ: Princeton University Press.

Mach E (1872, 1911) *History and Root of the Principle of the Conservation of Energy*, transl. Jourdain, P., Chicago: Open Court.

Machamer P, Darden L, Craver C (2000) Thinking about Mechanisms, *Philosophy of Science* 67(1): 1–25.

Morgan M S, Morrison M (eds) (1999) *Models as Mediators: Perspectives on Natural and Social Science.* Cambridge: Cambridge University Press.

Pitt J (2001) The Dilemma of Case Studies: Toward a Heraclitian Philosophy of Science, *Perspectives on Science* 9(4): 373–382.

Pritchard D (2009) Knowledge, Understanding and Epistemic Value. In *Epistemology, Royal Institute of Philosophy Supplement* 64, ed. O'Hear, A., 19–43, Cambridge: Cambridge University Press.

Psillos S (2017) World-involving Scientific Understanding. *Balkan Journal of Philosophy* 9(1): 5–18.

Railton P (1981) Probability, Explanation, and Information., *Synthese* 48(2), 233–256.

Salmon W (1998) *Causality and Explanation.* NY: Oxford University Press.

Scriven M (1965) Explanations, Predictions and Laws. In *Scientific Explanation, Space and Time*, eds H. Feigl, G. Maxwell, 170–230, Minneapolis: University of Minnesota Press.

Schickore J (2011) More Thoughts on HPS: Another 20 Years Later. *Perspectives on Science* 19(4): 453–481.

Strevens M (2008) *Depth: An Account of Scientific Explanation.* Cambridge, MA: Harvard University Press.

Trout J D (2002) Scientific Explanation and the Sense of Understanding. *Philosophy of Science* 69(2):198–208

Trout J D (2007) The Psychology of Scientific Explanation, *Philosophy Compass* 2:564–591.

Van Camp W (2014) Explaining Understanding (or Understanding Explanation). *European Journal for Philosophy of Science* 4(1): 95–114.

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

Woodward J (2003) *Making Things Happen: A Theory of Causal Explanation.* Oxford: Oxford University Press.

Zagzebski L (1996) *Virtues of the mind : an inquiry into the nature of virtue and the ethical foundations of knowledge.* NY: Cambridge University Press.