philosophical inquiries
www.philinq.it

philinq

philosophical inquiries

X, 1
2022

Edizioni ETS

Essays published on “Philosophical Inquiries” are double-blind peer-reviewed.

Please visit www.philinq.it for submissions and guidelines.

six-monthly journal / periodico semestrale
Subscription (paper, individual): Italy € 50,00, Abroad € 80,00
Subscription (paper, institution): Italy € 60,00, Abroad € 100,00

Subscription fee payable via Bank transfer to
Edizioni ETS
Intesa San Paolo
IBAN IT 21 U 03069 14010 100000001781
BIC/SWIFT BCITITMM
reason: abbonamento “Philosophical Inquiries”

Registrazione presso il Tribunale di Pisa n. 1/13

Direttore responsabile: Alessandra Borghini

EDIZIONI ETS
Palazzo Roncioni - Lungarno Mediceo, 16, I-56127 Pisa
info@edizioniets.com
www.edizioniets.com
Finito di stampare nel mese di agosto 2021

Distribuzione
Messaggerie Libri SPA
Sede legale: via G. Verdi 8 - 20090 Assago (MI)

Promozione
PDE PROMOZIONE SRL
via Zago 2/2 - 40128 Bologna

ISBN 978-884676420-1
ISSN 2281-8618

Table of contents

Essays

Focus

The conceptual framework of persons. A metaphilosophical investigation

philinq X, 1-2022
ISSN (print) 2281-8618-ETS

Past Present

Essays

A case for conservative ontology development in scientific metaphysics

Sahana V. Rajan

Abstract: Over the past decade, in contrast to the traditional analytic version of metaphysics, a brand of metaphysics that prioritizes collaboration and corroboration with sciences has emerged in the form of scientific metaphysics. While there has been a shift from the methodological dependence of analytic metaphysis on intuition, and conceptual analysis to the methodological preference for empirically-motivated metaphysical insights in scientific metaphysics, such a shift has not penetrated the foundational aims. Scientific metaphysics continues to probe the nature and structure of reality, much like its analytic counterpart and in this process, develops ontologies. Broadly two kinds of ontologies are furnished - global metaphysical ontologies and local scientific ontologies. In this paper, I highlight the challenges with developing such ontologies in scientific metaphysics. With Ladyman-Ross' Information Theoretic Structural Realism as a case in point, I contest that the former suffers from representational indeterminacy and redundancy. Further, I note the possibility that eventually, local scientific ontologies might be replaced by scientific theories and in such a scenario, the former are best conceived as interim metaphysical supports for the latter.

Keywords: scientific metaphysics; ontologies; representation; redundancy.

> A metaphysical philosophy, in the sense of that which is to be definitively accepted in advance of scientific inquiry, is, or should be, a system of pigeon holes in which facts are to be filed away. Its first merit is to give a place to every possible fact. Whatever could conceivably be settled by experiment, metaphysics should abstain from settling in advance.
>
> (Peirce 1975: 201)

1. *Introduction*

The pigeon holes are ready. A hundred ontological systems, and a thousand eager metaphysicians await. Where are the scientific facts? Last few centuries,

philinq X, 1-2022, pp. 9-24
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.295
Submitted: May 2020
Accepted: February 2021

many scientific facts arrived (think of phlogiston theory, caloric theory, or Newtonian physics) and were safely filed away in the pigeon holes. But soon after, these 'facts' retired to the overcrowded hall with other superseded scientific theories and the pigeon holes were emptied. The latest facts (think of Quantum Field Theory, Modern Evolutionary Synthesis) to occupy these pigeon holes have survived a great many tests- how long will the current tenants last[1]?

Metaphysicians in the enterprise of gathering and filing away scientific facts into ontological pigeon holes should be ubiquitously aware of the possibility that such facts are always prone to eviction by the next drove of facts. The acceptance of scientific theories is to be solemnly accepted with the possibility of their future retirement. Such possibilities have invoked a cautious optimism in scientific realists and pessimism in scientific anti-realists about the status of scientific theories. Scientific realists defend their optimism by citing the explanatory and predictive successes of the retired theories, in form of the no-miracles argument (Brock and Mares 2007; Devitt 2008; Park 2019). For now, in the course of this paper, I side with the hesitant optimism of a scientific realist, supported by empirical evidence which indicates that scientific theories can latch onto and mathematically specify the structures of our reality (Worrall 1989; Ladyman 1998; Arenhart and Bueno 2015). However, ontologies have not always been built, in the history of metaphysics, in such strict corroboration with scientific theories. The coming together of science and metaphysics is, in many ways, a recent phenomenon. The traditional manners of building ontologies often relied on *a priori* methodologies, including intuition and conceptual analysis. Such ontologies have been critiqued for misrepresenting, or in Ladyman's terms, domesticating science to advance their metaphysical points (Ladyman *et al* 2007: 4-8)[2]. In contrast to the reliance on

[1] Acknowledgments: I am indebted to the support and guidance of two generous anonymous reviewers. I am grateful for their help throughout the process of developing and formulating this paper. Many thanks to Aditya Jha (University of Canterbury), Alex Franklin (King's College London), Ana-Maria Cretu (University of Bristol), Anjan Chakravartty (University of Miami), Joaquim Giannotti (University of Glasgow), Mahmood Vahidnia (Shahid Beheshti University) and Michela Massimi (University of Edinburgh) whose comments on my presentation at the Methodological Issues in Metaphysics of Science Workshop 2020 have immensely helped me in formulating the insights examined in this paper.

[2] The representation of analytic metaphysics by James Ladyman, Don Ross, David Spurrett and others (hereon, RLS) has been severely challenged over the past decade. The dangers of sweeping generalizations, possible misrepresentation and the absence of in-depth examination of neo-Scholastic metaphysical positions as well as doubts about unified worldview and nature of a scientistic metaphysics were brought forth by P. Kyle Sanford, Katherine Hawley, Paul Humphreys, Cian Dorr among others (Stanford *et al.* 2010; Dorr 2010). Although the discussion of these challenges is crucial, it is unfortunately beyond the scope of my paper. While receiving RLS's comments on analytic metaphysics, the reader should note the limitations of their accounts and also note the extensive ongoing debates on methodologies in metaphysics (Thomasson 2007; 2012; Nolan 2016; Lee 2017).

a priori methodologies, over the past decade, there has been a concerted effort to develop Scientific Metaphysics[3] that endorses an empirically-motivated form of metaphysical theorising in collaboration and corroboration with scientific theories (Ross *et al.* 2013; Soto 2015).

While developing ontologies in analytic metaphysics, agreement with scientific facts is an optional feature. This has often resulted in metaphysical hypotheses that are divorced from scientific findings about the world and have been the target of glaring critique by many metaphysicians (Ladyman and Ross 2007: ch. 1; Maclaurin *et al.* 2012; Byrant 2020). In contrast, such an agreement is a necessary (though, not sufficient) condition for theories to qualify as naturalised or scientific, in case of naturalised or scientific metaphysics. Under such an agreement to scientific theories, scientific metaphysics can develop, I propose, two broad kinds of ontologies, global metaphysical ontologies and local scientific ontologies. I contest that the former suffer from representational indeterminacy and redundancy. With Information Theoretic Structural Realism (ITSR) of RLS[4] as a case in point, I suggest that it would be a wise move to eliminate global ontologies. Further, as scientific theories develop, the underdetermination of metaphysical claims and local ontologies could be gently addressed. In the end, I note the possibility that eventually, local scientific ontologies might be replaced by scientific theories and in such a scenario, the former are best conceived as interim metaphysical supports for the latter.

To this end, here is a roadmap of what follows: I begin by introducing The Reader's House which eases us into distinction between local scientific ontologies and global metaphysical ontologies in Section 2. The elimination of global ontologies is proposed with ITSR as a case in point in Section 3. Further, the underdetermination of local scientific ontologies by scientific theories is discussed in Section 4, highlighting the possibility of the elimination of those ontologies that do not accommodate scientific developments and also, the eventual collapse of the surviving local scientific ontology onto scientific theories. In the end, I briefly note the profit of a conservative ontology development for scientific metaphysics (Section 5).

[3] Scientific metaphysics differs from metaphysics of science in some crucial ways. Scientific Metaphysics is a specific form of metaphysical theorising committed to collaboration and corroboration with scientific theories; it is a counteraction to the dominance of *a priori* methodologies in traditional analytic metaphysics. Metaphysics of science involves (but is not exclusive to) debates relating to metaphysical aspects of various scientific domains (physics, chemistry, biology, psychology, economics and others). There can be both analytic and scientific metaphysics of science. In case of the former, *a priori* methodologies can be utilized to infer metaphysical claims based on scientific theories. However, scientific metaphysics prioritizes scientifically-informed metaphysical claims.

[4] This is a short form for "Don Ross, James Ladyman, David Spurrett" and is borrowed from Melnyk 2013.

2. *Conservative ontology development: a starter's kit*

2.1. The reader's house

Imagine a metaphysical theory called "The Reader's House" which maps the ontology of your house. You are the metaphysician developing this comprehensive inventory. As you walk through your house, you make a numbered list: you write down the name of every item in your house starting with the doorbell and name plate to the fence in your backyard. Through this process, depending on your preference for detail, you might want to go a little deeper and even add the length of wires used in the electrical apparatus running in your house; perhaps, you feel a little adventurous and also put down the screws, nuts and bolts. Right now, you have a numbered list of all items in your house. As you go through this giant list, you realize that you have left a precious load of items: what about everything that *constitutes* all of these things in your house? You wonder: What about the quarks, the electrons, the atoms, the molecules and everything between and within? And this triggers another rabbit hole of ontological meditations: what about your dogs? What about your partner and the kids? And then: What about the ways in which each of them relates to each of the other and also, how each of us relate to the humongous list of items you jotted down? Unless you pin down all of this, you would be missing out on accurately representing the ontology of your house. This leads to an inventory which takes the form of an infinitely cross-referenced complex network. This entire enterprise assumes you have collaboration *par excellence* with physicists, chemists, psychologists, sociologists and others who enable you to correctly classify everything.

You sit with the colossal colour-highlighted, neatly bound database and think: "Something must be missing!", a feeling that haunts most metaphysicians even after stuffing their ontologies as much as possible. You rest a bit, watching your dog, Maya. Looking at her, imagining so many ways in which we all are common, you wonder: "Could there just be one broad category, to which all of us could belong? Could I headline the inventory such that everything falls under it?". You think through the options: there could be that-which-is, material beings or physical entities, perhaps?

This movement marks a paradigmatic shift. You are not talking about specific kinds of things anymore, which are studied by scientific theories- you are moving to make general claims about everything, all that is. This is the jump from ontological claims which can be strictly warranted by your squad of scientific experts to metaphysical claims which do not fall under any of their specialisations. You are the creator here. This is where you, the metaphysician, run the risk of indulging in the excesses of the *a priori*.

If you have so far been committed to scientific metaphysics in developing The Reader's House, I imagine that you corroborated your claims with your group of scientific experts. However, as you distance and intend to widen the scope of your claims, you take on the perspective of a vantage point which no singular science could offer you. Till this point, you depend on the rigor that the individual scientific theories supply to your inventory- hereon, you are on your own, building a panoptic perspective. In the coming section, I introduce the notions of local and global ontologies to talk about this movement from ontologies supported by scientific theories to ontologies which move over and beyond them.

2.2. Local scientific ontologies and global metaphysical ontology

As you gradually extended The Reader's House inventory, at some point, a pragmatically motivated project of developing a list of items in the house turned into a larger metaphysical project, intended to capture all-there-is. In both of these projects, the larger goal is to represent the world around you. While developing an inventory of list of items in the house, you represent a specific set of phenomena. In corroboration with the relevant scientific theories, you jot down the entities and processes which are empirically supported. Ignoring, for now, the deeper details about the timeline and order in which you noted the items, you covered the following phenomena amongst many others. There are the quantum phenomena, then there is virus, chromosomes, red blood cells and then come the pollens. Then came amoeba, ticks, dust mites so on and on until you reached apples, a mouse, a soccer ball and your dog Maya and eventually, you measured the house itself. In each of these cases, a specific set of phenomena is investigated, and you depend upon the relevant scientific theory to offer empirically supported information about their nature or structure and their behaviour. Metaphysical claims and ontologies specific to these set of phenomena can be proposed, in collaboration with the scientific theories. Local scientific ontologies are those ontologies that are developed and based on particular theories, and are restricted to a scientific domain; they are theory and domain restrictive. Consider neuroscience which broadly studies the nervous system, more specifically the brain and its structure and development, its effect on our cognitive functions and behaviour. Based on the findings and developments in neuroscience, we can arrive at a multitude of local scientific ontologies (like Integrated Information Theory, Global Workspace Theory, Recurrent Processing Theory) which attempt to metaphysically capture the behaviour of certain kinds of organisms. In these local scientific ontologies, any metaphysical claim receives its confirmation or falsification from the findings and developments in neuroscience. This is, however, not the case

with building an ontology that fits reality or the world. Global metaphysical ontologies are those ontologies developed to represent any phenomena whatsoever. Such ontologies are not based on particular scientific theories or on empirical information pertaining to a scientific domain; they are domain and theory neutral. In the coming section, I suggest that global ontologies suffer from representational indeterminacy and redundancy, which motivates their elimination in scientific metaphysics[5].

3. *Decluttering the reader's house: elimination of indeterminate and redundant global ontologies*

In scientific metaphysics, a global ontology posits entities and processes (like properties, universals, hunks of matter or real patterns) that underlie or are presupposed by scientific theories. The world is believed to possess a fundamental nature or fundamental structure, constituted by such entities and processes. Existence claims of global ontology take forms such as, "It is real patterns all the way down." (Ladyman and Ross 2007: 228) or for analytic metaphysics, "I propose that a physical object is not an enduring spatial hunk of matter, but is, rather, a spatiotemporal hunk of matter." (Heller 2001: 331). Such global ontologies which aim to represent 'world' or 'reality' suffer from representational indeterminacy.

3.1. The problem of representational indeterminacy

Consider a global ontology $O_{G.1}$ which declares that the world is constituted by substances possessing essential and accidental properties as well as internal and external relations. To the question, "What phenomena does $O_{G.1}$ represent?", we could expect the response, "*Everything* around us.". If pushed further, "What is *everything*?", $O_{G.1}$ could come back with, "You can pick

[5] The proposed distinction between global metaphysical and local scientific ontologies should be differentiated from the classification of globally applied and locally applied naturalistic metaphysics suggested by Soto 2017. Soto considers globally applied naturalistic metaphysics to be a component of metaphysical practice that investigates ontological issues relating to fundamental structure of reality (i.e. those features of our physical world that can be instantiated everywhere in the world). These would include questions about whether space and time are relational or substantival or whether reality has a natural-kind structure. Also, locally applied naturalistic metaphysic are specific debates which arise pertaining to the role of unobservable posits within particular theories of scientific ontology. He illustrates this through the case of dark matter. In my proposal, global metaphysical ontologies make claims about the 'world' or 'reality' and I suggest their elimination due to indeterminacy and redundancy. Specifically, questions about fundamental structure or nature, a natural-kind structure, or the role of unobservables sans the context of specific sciences would be treated with sceptical hesitation and might face elimination eventually. Local scientific ontologies, in my proposal, are those ontologies which are based on scientific theories and are underdetermined by the latter.

up *anything* that is happening *anywhere*." In reply, a complaint to the effect, "That is vague. What are the specific phenomena that $O_{G.1}$ attempt to capture?" might extort a mildly concrete response, "The ontology can offer the metaphysical nature of, say, this mountain, that ball or even your dog.". Global ontologies suffer from representational indeterminacy, the lack of targets of representation and require probing to clarify the determinate targets. Once such targets are identified, scientific metaphysicians can refer to relevant sciences which would provide empirical information and form the context within which the metaphysical claims and entities and processes posited by $O_{G.1}$ can be construed. In the above case, a scientific metaphysician could cite mountain geography to learn about the mountain, physics to study more about the ball and evolutionary biology and zoology to find out more about my dog. It might be fitting to imagine $O_{G.1}$ as an abstract world of Platonic forms which awaits its instantiation in the sciences. Such abstract worlds, characterised by representational indeterminacy, are neither empirically-motivated nor scientifically-informed and should be eliminated.

Some metaphysicians of science might remind us that it is not possible to completely cleanse a scientific theory of *a priori* claims, since theories inevitably *presuppose* metaphysical details (Chakravartty 2010; 2013, Mumford and Tugby 2013). However, such metaphysical matters can be fruitfully analysed within the bounds of scientific theories and need not be generalized to the degree of indeterminacy. Ladyman also seems to support such restriction of metaphysical theorising within scientific contexts, when he emphasises that the assumption of a general composition relation, beyond "the particular kinds of composition relevant to their respective domains" is symptomatic of nothing more than "an entrenched philosophical fetish" (Ladyman and Ross 2007: 21, Ladyman 2012). In the coming section, I highlight that scientifically-informed global ontologies, in addition to representational indeterminacy, also suffer from redundancy.

3.2. The Redundancy Dilemma

The Redundancy Dilemma applies to those global ontologies, that offer an inventory of entities and processes for the 'world', based on scientific theories. Here is a prototype: Suppose a global ontology $O_{G.2}$ that says that the world is constituted by structures, that further consist of relations between phenomena. $O_{G.2}$ is based a scientific theory T_S at time *t*. T_S gradually develops over decades. Assuming that it is not proven false or superseded, imagine time $t+_n$ when T_S can reductively or non-reductively represent a range of phenomena (including non-physical). On such an exciting and promising day, T_S would represent the world in terms of entities and processes that empirically constitute its theory.

An example could be the case of Quantum Field Theory which mathematically represents the world in terms of quantum fields, particles, waves (Kuhlmann *et al.* 2002, Baker 2016). In a more developed form, QFT would continue to represent the world in terms of these ontological categories. In such a scenario, $O_{G.2}$ would be redundant. The categories of structures and relations could be powerful conceptual instruments *in* scientific theories. However, there is no good reason to assume there ought to be something to represent beyond the determinate representational targets of scientific theories. More importantly, reiterating the point that Ladyman raised earlier, global ontologies forgo intricate representational details of scientific theories. In practice, representation of each phenomenon is a complex function of the pre-existent scientific discourse on the phenomena, the pragmatics of technological scope and limitations and other factors (Coopmans *et al* 2014). The Big Picture, the grand narrative of the 'world' or 'reality', goes beyond the determinate targets of representation, and there seems to be no good reason to believe in such a narrative.

Global ontologies are representationally indeterminate and redundant. Their elimination would regulate the conceptual gymnastics performed in scientific metaphysics based on purely *a priori* methodologies. In the next section, I illustrate the case for their elimination through Information-Theoretic Structural Realism of James Ladyman, Don Ross, David Spurrett and others.

3.3. Information Theoretic Structural Realism of Ladyman, Ross and Spurrett

In the previous section, I questioned The Big Picture generally presupposed by a global ontology. An instance of such a global ontology is Information Theoretic Structural Realism (ITSR).

From a scientistic stance (synthesis of empiricist and materialist commitments), RLS adopt the Dennettian theory of real patterns and redevelop it into Information Theoretic Structural Realism (ITSR). This is a global ontology for a non-reductive unified worldview, which proposes that our world is constituted by representational and extra-representational or universal real patterns (Dennett 1991).

RLS define real patterns in the following way:

> A pattern P is real iff
> (i) it is projectible; and
> (ii) it has a model that carries information about at least one pattern D in an encoding that has logical depth less than the bit-map encoding of D, and where D is not projectible by a physically possible device computing information about another real pattern of lower logical depth than P. (Ladyman and Ross 2007: 233)

Essentially, a pattern is a regularity in data (which can be understood, roughly, as that which is observed or is observable). Some such regularities in data can be compressed through relevant algorithms. Depending on its computational capacities, an observer compresses continuous flow of such data and produces a model which offers a better-than-chance prediction of future events. These processes, of observation (in relation to capacities of the physically possible machine), compression and processing of data as well as the production of predictive models constitutes a real pattern (Ladyman and Ross 2013). There are largely two kinds of real patterns, representational real patterns and extra-representational or universal real patterns.

On one hand, representational real patterns are: (i) Predictive models produced by compression of observed data, relative to the computational capacities of the observers; (ii) Such models are expressed through natural or ordinary language (employing notional-world concepts like causation and cohesion). Such predictive models capture ontologies, relative to specific scales of measurement, at which observation occurs. RLS suggest that special sciences (like non-fundamental branches of physics, chemistry, biology, economics, sociology) trade in such representational real patterns. Metaphysical claims of special sciences are largely epistemological products which enable measurement conducive to the computational convenience of observers of a certain kind; a case in point is the category of individuals which is considered to be an epistemological book-keeping device.

On the other hand, extra-representational or universal real patterns are those predictive models, mathematically specified by fundamental physics, which capture the entire physical complexity of the phenomenon. In such patterns, there is lossless compression of observed data. Fundamental physics generates such patterns and represents the objective modal structure of our world. Such structures can be measured at any scale of measurement (i.e. they are scale-neutral and thus, are "universal" real patterns). RLS admit that the goal of scale-neutral real patterns capturing the complexity of the world in its entirety is a limiting ideal, an aspiration which perhaps, we would never reach[6].

[6] The claim that we might never reach such an ideal is insufficient to insulate ITSR from the redundancy dilemma. If we might never reach such an ideal, what purpose does it effectively serve for scientific metaphysics? The two regulative principles proposed by RLS, namely, Principle of Naturalistic Closure (PNC) and Primacy of Physics Constraint (PPC) can limit and moderate metaphysical claims in Ontic Structural Realism (Ladyman *et al.* 2007: 37-38). The aspiration for a completed fundamental physics might be reflective of the subjective preferences of the theorists. Such preferences are undeniably an important aspect of the epistemic factors which govern the development of scientific theories and the derivative local ontologies. However, when taken out of the context of scientific theories and local ontologies, such aspirations are often symptomatic of, in Ladyman's terms, "an entrenched philosophical fetish" for The Big Picture.

Illustrating through Eddington's famous instance of two tables, RLS explicate the relation between universal real patterns of fundamental physics and representational real patterns of special sciences (Ladyman 2018). The table of everyday life is a real pattern projectible from a specific macroscopic scale. However, at a fundamentally physical scale, the table is a pattern of molecules which are attached together by electromagnetic potentials. We can, at best, correlate the everyday-life-table as a representational real pattern to the fundamentally-physical bound state of particles which composes it[7].

ITSR is a global ontology of real patterns for a unified non-reductionistic worldview. This unity is specified in the mathematical structures of fundamental physics, which captures the objective modal structure of the world. In the next section, I show the redundancy of such an ontology and suggest its elimination.

3.4. ITSR: eliminating The Big Picture

Suppose that today is the day that fundamental physics can universally measure real patterns and reductively or non-reductively represent different kinds of phenomena. In such a scenario, those entities and processes that empirically constitute the fundamentally physical theory are measured and represented. To illustrate, say, QFT takes the form of such a fundamental physics, then we would measure and represent determinate targets including particles, waves, fields. The ontological categories of representational or universal real patterns would neither play any role nor add no metaphysical value to the scientific theory. In other words, ITSR would be redundant.

While we wait for fundamental physics to develop further, we could take inspiration from RLS who say the following of scientists pursuing ultimate answers to big metaphysical questions: "Scientists who rush to pronounce on such questions in the light of the latest theories go beyond the evidence." (Ladyman and Ross 2013: 131). ITSR is an instance of "...going beyond the evidence" and while theories develop, it might be advisable to avoid rushing to global ontologies.

A disunity of science zealot or an admirer of promiscuous realism might regard the suggestion for contextualisation of metaphysical claims within diverse sciences as an affirmative for a disunified worldview. However, the thesis

[7] The two tables cannot be identified with one another because they possess different persistence and modal properties. While the table could exist even after some of its relevant particles did not, this does not stand for the bound state which would alter with the change of these particles. Also, though the table could have a leg replaced, the bound state would not survive such a shift. For a detailed examination of ITSRist relationship between fundamental physics and special sciences, see Ladyman 2009.

of disunity of sciences is as excessive as the thesis of unity of sciences. Both assume a substantive notion of a metaphysical reality, so to say, The Big Picture of our world that is either one or broken, which is either unitary or fragmented. The recommendation of global ontology eliminativism declares that there is no Big Picture of the world. The theses of unified and disunified world-views are equally at fault for assuming there is such a picture, a world beyond the determinate targets of representation.

In this section, the eliminativist stance towards global ontologies has been discussed. Based on the instance of ITSR, I have emphasized that a developed form of fundamental physics could render ITSR redundant. In the next section, I consider underdetermination of local ontologies derived from or based on scientific theories. Such local ontologies would face elimination as scientific theories develop. Those ontologies that can accommodate scientific developments would survive while the ones which fail to account for them would perish.

4. *Local scientific ontologies*

4.1. Indeterminacy of general metaphysical notions

Let's go back to the pigeon holes of metaphysics. A local scientific ontology constitutes pigeon holes to stack away the facts of a domain-specific current best scientific theory. Say, to accommodate QFT of fundamental physics, Ontic Structural Realism suggests the pigeon holes of structures, relations and interactions. Consider Integrated Information Theory that suggests the pigeon holes of systems with parts involved in cause-effect feedback loops for accommodating findings of affective, cognitive and computational neurosciences. The previous argument against global ontology underscored that there is no reason to assume that there is a general set of pigeon holes which can accommodate all scientific theories. Unfortunately, a general set of pigeon holes cannot even be cashed out for a domain-specific scientific theory. From a scientific theory, a multitude of local scientific ontologies can be derived. Global Workspace Theory, Integrated Information Theory, and Recurrent Processing Theory are some of the local scientific ontologies based on neuroscience. From Quantum Field Theory of fundamental physics, we can derive the traditional substance-attribute metaphysic, ontology of individuals and sets, ontologies of facts, Whiteheadian process ontology, ontology of tropes and trope bundles as well as field and structuralist ontologies (Kuhlmann *et al.* 2002). To this morass of ontologies is also added another can of worms, metaphysical claims that hint at seemingly opposing views. A recent example is the debate on quantum objects as individuals and as non-individuals (French and Krause 2006:

Chapter 4, Caulton 2015, Arenhart 2015). This signals the underdetermination of local ontologies and metaphysical claims by scientific theories (Laudan and Leplin 1991, Stanford 2001, Norton 2008, Magnus 2010, Werndl 2013). The Big Picture is, then, this: A rainforest of local ontologies flooded with often-incompatible metaphysical claims.

A surveyor of this rainforest might observe common categories (like structures, substances, individuals, properties) and intuitively infer a global ontology constituted by these categories. However, this would be an unwarranted jump. There are significant scientific details in a theory which support the choice of an ontological category within local ontologies and a generalisation to the effect that structures (or any other ontological category) constitute our world turns a blind eye to important aspects of the scientific theories within which the notion of structure figures. In QFT, the non-individuality of particles motivates the category of structures. In neuroscience, an evolutionary trajectory of the brain as an organ and specifically, the circuitry of brain regions and activation of neural networks might motivate the category of complex systems or structures. In each of these cases, there are crucial differences in the scientific details that underlie the choice of category. The way in which neural activations occur and form dynamic brain networks is significantly different from the way in which particles behave at a subatomic level. There is no reason to assume a substantive notion of structure to denote the phenomena across domains or theories. A safer bet to acknowledge such commonalities across ontologies could be to suggest that there are Suárezian surface features to the notion of structure, which characterises its use in diverse scientific theories (Suárez 2010; 2015). That a structure is constituted by relations between phenomena could be its surface feature, capturing the most general feature that marks its use across theories. However, the utility of general notions and surface features for scientific theories is negligible. To put such notions to use, the metaphysician ought to engage with un-domesticated versions of scientific theories and recognize the relevant instantiations of structures and relations within the theory. If the use of a metaphysical claim or ontological category requires acquaintance with the relevant scientific theories, then what could be the role of general notions? Here, the Redundancy Dilemma resurfaces.

4.2. The end fame for local scientific ontologies: victorious replacement by scientific theories

As scientific theories develop, the underdetermination of local scientific ontologies could gradually resolve. Lesser and lesser number of local ontologies would continue to accommodate the latest developments in scientific theories. As neuroscience developed, Cartesian Dualism, a local ontology based on

anatomical and physiological studies undertaken by Descartes, was abandoned and replaced by other theories of mind or consciousness (such as Identity Theory, Functionalism, Behaviourism, Eliminative Materialism) that acknowledge the role of the brain in the identity or emergence of mental states.

With gradual elimination of local ontologies, scientific metaphysicians might realise that the ultimate victory for a local scientific ontology is to collapse onto a scientific theory. In neuroscience, Eliminative Materialism is a local ontology that predicts that the endgame for theories of mind and consciousness is their elimination and replacement by a complete neuroscience (Churchland 1981). Scientific metaphysicians ought to seriously consider the possibility that local scientific ontologies might be, at best, interim metaphysical devices that serve to advance scientific theories and whose success is the extent to which they can collapse onto developed versions of the theories.

5. *The case for conservative ontology development*

Scientific metaphysics began as an attempt to regulate metaphysical theorising, previously governed solely by intuition and conceptual analysis. This regulation was carried out by committing to rigorous engagement with scientific theories. Over the past decade, in metaphysical debates that value collaboration and corroboration with scientific theories, there have been signs that intuition and conceptual analysis could override the strict boundaries drawn by the latter; this is exemplified by the case of global ontologies. Global ontologies are representationally indeterminate and redundant and should be eliminated. As a metaphysically healthy alternative, we can develop a range of local ontologies based on scientific theories. Of these local ontologies, only those which accommodate scientific developments would continue to be relevant and eventually, might face a collapse onto developed versions of scientific theories.

To conclude, I restate the case for Conservative Ontology Development: first of all, global ontologies that declare an ontology for all-there-is, world, or reality suffer from representational indeterminacy and redundancy and should be eliminated, and second, the relevance of local scientific ontologies is determined by the extent to which they can accommodate the latest scientific developments. With such conservative ontology development, scientific metaphysics can avoid excessive dependence on *a priori* and can channelize its theoretical resources towards serving scientific theories to reach more developed forms.

Sahana V Rajan
Jindal Global Business School, O P Jindal Global University
svrajan@jgu.edu.in

References

Arenhart, Jonas R. B., 2017, "The received view on quantum non-individuality: formal and metaphysical analysis", in *Synthese*, 194: 1323-1347.

Arenhart, Jonas R.B., and Otávio Bueno, 2015, "Structural realism and the nature of structure", in *European Journal for Philosophy of Science*, 5: 111-139.

Baker, David John, 2016, "The Philosophy of Quantum Field Theory", in *Oxford Handbooks Online*, Oxford University Press.

Byrant, Amanda, 2020, "Keep the chickens cooped: the epistemic inadequacy of free-range metaphysics", in *Synthese*, 197: 1867-1887.

Brock, Stuart and Edwin Mares, 2007, *Realism and Anti-Realism*, Acumen Publishers, Durham.

Caulton, Adam, 2015, *Issues of identity and individuality in quantum mechanics* (Doctoral thesis), University of Cambridge, <https://www.repository.cam.ac.uk/handle/1810/249007>.

Chakravartty, Anjan, 2010, "Metaphysics between the Sciences and Philosophies of Science", in P.D. Magnus, J. Busch, eds., *New Waves in Philosophy of Science*, Palgrave MacMillan, London.

Chakravartty, Anjan, 2013, "On the Prospects of Naturalized Metaphysics" in Don Ross, James Ladyman, and Harold Kincaid, eds., *Scientific Metaphysics*, Oxford University Press, Oxford, UK: 27-51.

Churchland, Paul, 1981, "Eliminative Materialism and Propositional Attitudes", in *The Journal of Philosophy*, 78, 2: 67-90.

Coopmans, Catelijne, Janet Vertesi, Michael E. Lynch, and Steve Woolgar, eds*.,* 2014, *Representation in Scientific Practice Revisited,* MIT Press, Cambridge.

Dennett, Daniel, 1991, "Real Patterns", in *The Journal of Philosophy*, 88, 1: 27-51.

Devitt, Michael, 2008, "Realism/Anti-Realism", in Stathis Psillos and Martin Curd, eds., *The Routledge Companion to Philosophy of Science*, Routledge, London.

Dorr, Cian, 2010, "Review of James Ladyman and Don Ross, Every Thing Must Go: Metaphysics Naturalized", in *Notre Dame Philosophical Reviews* 6, <https://ndpr.nd.edu/news/every-thing-must-go-metaphysics-naturalized/>.

French, Steven, Décio Krause, 2006, *Identity in Physics: A Historical, Philosophical, and Formal Analysis*, Oxford University Press, Oxford.

Heller, Mark, 2001, "Temporal Parts of Four-Dimensional Objects" in Michael J. Loux (ed.), *Metaphysics: Contemporary Readings*, Psychology Press, London.

Kuhlmann, Meinard, Holger Lyre, and Andrew Wayne, 2002, *Ontological Aspects of Quantum Field Theory*, World Scientific Press, Singapore.

Ladyman, James, 1998, "What is structural realism?", in *Studies in History and Philosophy of Science Part A*, 29, 3: 409-424.

Ladyman, James and Don Ross, 2007, *Every Thing Must Go: Metaphysics Naturalized*, Oxford University Press, Oxford.

Ladyman, James, 2009, "Weak Physicalism and Special Science Ontology", in Alexander Hieke and Hanne Leitgeb, eds., *Reduction, Abstraction, Analysis,* Ontos Verlag, Germany.

Ladyman, James, 2012, "Science, Metaphysics and Method", in *Philosophical Studies,* 160: 31-51

Ladyman, James, 2018, "Scientific Realism Again", in *Spontaneous Generations: A Journal for the History and Philosophy of Science*, 9, 1: 99-107.

Ladyman, James, Don Ross, 2013, "The World in the Data", in Don Ross, James Ladyman, and Harold Kincaid, eds., *Scientific Metaphysics*, Oxford University Press, Oxford, UK: 108-150.

Laudan, Larry, Jarrett Leplin, 1991, "Empirical Equivalence and Underdetermination", in *Journal of Philosophy*, 88: 449-72.

Maclaurin, James, Heather Dyke, 2012, "What is analytic metaphysics for?", in *Australasian Journal of Philosophy*, 90, 2: 291-306.

Magnus, P.D., 2010, "Inductions, Red Herrings, and the Best Explanation for the Mixed Record of Science", in *British Journal for the Philosophy of Science*, 61, 4: 803-819.

Melnyk, Andrew, 2013, "Can Metaphysics Be Naturalized? And If So, How?" in Don Ross, James Ladyman, Harold Kincaid, eds., *Scientific Metaphysics*, Oxford University Press, Oxford, UK: 79-95.

Mumford, Stephen, and Matthew Tugby, 2013, "What is the Metaphysics of Science?" in S. Mumford and M. Tugby, eds., *Metaphysics and Science*, Oxford University Press, Oxford.

Nolan, Daniel, 2016, "Method in Analytic Metaphysics", in *The Oxford Handbook of Philosophical Methodology*, Herman Cappelen, Tamar Szabo Gendler, John Hawthrone, eds., Oxford University Press, New York.

Norton, J, 2008, "Must Evidence Underdetermine Theory?", in M. Carrier, D. Howard, and J. Kourany, eds., *The Challenge of the Social and the Pressure of Practice: Science and Values Revisited*, University of Pittsburgh Press, Pittsburgh.

Park, Seungbae, 2019, "How to Formulate Scientific Realism and Antirealism", in *Journal for General Philosophy of Science*, 50, 4: 477-488.

Peirce, Charles Sanders, 1975, *Charles Sanders Peirce: Contributions to The Nation. Volume 1*, ed. by K.L. Ketner and J. Cook, Texas Tech University Press, Cambridge, Lubbock.

Soto, Cristian, 2015, "The Current State of The Metaphysics of Science Debate", in *Philosophica*, 90, 23-60.

Soto, Cristian, 2017, "Globally and locally applied naturalistic metaphysics", in *Manuscrito*, 90, 3: 33-50.

Stanford, P. Kyle, 2001, "Refusing the Devil's Bargain: What Kind of Underdetermination Should We Take Seriously?", in *Philosophy of Science*, 68, S3: s1-s12

Stanford, P. Kyle, Paul Humphreys, Katherine Hawley, James Ladyman, Don Ross, 2010, "Protecting rainforest realism", in *Metascience* 19: 161-185.

Suárez, Mauricio, 2010, "Scientific Representation", in *Philosophy Compass*, 5(1): 91-101.

Thomasson, Amie L, 2007, "The Methods of Metaphysics", in *Ordinary Objects,* Oxford University Press, Oxford.

Thomasson, Amie L, 2012, "Research Problems and Methods in Metaphysics", in *The Continuum Companion to Metaphysics*, Robert Barnard and Neil Manson, eds., Continuum International, New York.

Werndl, Charlotte, 2013, "On Choosing Between Deterministic and Indeterministic Models: Underdetermination and Indirect Evidence", in *Synthese*, 190: 2243-2265.

Worrall, John, 1989, "Structural Realism: The Best of Both Worlds?", in *Dialectica*, 43, 1: 99-124.

I am hungry, therefore I am. Paul Ricœur's hermeneutic phenomenology as a model for food existentialism

Maria Cristina Clorinda Vendra

Abstract: By focusing on the hunger drive and the act of eating as existential dimensions, this essay considers the possibility to extend Paul Ricœur's thought in the direction of food philosophy. By conceiving his hermeneutic phenomenology as a model for food existentialism, this paper aims to discuss hunger and eating as interrelated aspects of human beings' embodied existence that are involved in the social world. I will begin with a phenomenological description of hunger and eating referring to Ricœur's analysis of the corporeal involuntary as offering the base features to develop what I will call an "interpretive existential philosophy of being hungry and eating". Then, I will turn to hunger and eating as involved in the real complexity of temporal experience. These reflections will lead to examining the interplay of cosmic time and lived time in relation to hunger and eating, opening up the discussion of the gustatory time through the intersection of the objective time of the clock and the subjective time of the stomach.

Keywords: hunger; eating; food philosophy; embodiment; gustatory time.

Hunger is a primary mode of being related to a complex array of neurophysiological states. On the one hand, in its most general sense hunger is an involuntary phenomena correlated with the act of eating. On the other hand, though, human hunger is far more complex than the hunger of other animals. Far from being just an automatic bodily mechanism, hunger goes beyond our senses, bodies, and brains. We train, resist, and stimulate hunger, that is, we manage it as part of our education (see Borghini 2016). Therefore, human being's relationship to nourishment cannot be reduced to a matter of pure sustainment. Representing the most complex ecological relationship in which we take part, as well as the most fundamental origin of the encounter between ourselves and the otherness of the world, hunger is a complex natural, cultural, and social aspect of human existence that call for interdisciplinary analyses. The recognition of the centrality of hunger and eating to the understanding of our existence results in the current explosion of food studies in many disciplines, among them those of the natural and social sciences, humanities and

philinq X, 1-2022, pp. 25-40
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.351
Submitted: December 2020
Accepted: October 2021

cultural theories. The possibility to produce critical knowledge in relation with hunger and eating as topics of high theoretical and practical relevance, has opened up the space for the development of the field of food philosophy. The question "how are we to eat?" (Boisvert *et al.* 2016: 7) arises in today's food-oriented philosophical thinking as one of the deepest problems a food philosopher could analyze. This essential query involves different meaning shifts ranging from health issues to cultural appropriateness. Profoundly affecting human being's access to the meanings and values belonging to different spheres of human life, such as the practical, nutritional, ethical, aesthetical, physical, ecological, epistemological and ontological sphere, hunger and eating undoubtedly deserve a thorough philosophical investigation.

The analysis of hunger and eating as hot topics in the philosophical field is a fairly recent phenomena. The reason for the increasing number of contributions to the branch of food philosophy can be briefly explained as follows. Whereas in the modern philosophical inquiry hunger and eating were seen as aspects connected to human being's instincts and bodily condition considered as the lowest degrees of human existence, the attention given to the lived body in twentieth-century philosophy, which invites us to consider the profound importance of embodiment in how we think and act, enables philosophers to focus on hunger and eating as significant themes in the current philosophical debate. The renewed attentiveness to the phenomenon of human corporeality makes hunger and eating arise as complex conditions of human existence. The specific meaning of hunger and eating for the human being cannot be, then, understood without paying attention to our embodied existence and to our interdependence with the social environment as essential surrounding and lived space in which our life takes place together with all other animate species and inanimate things (see Vendra 2020).

It is from the consideration of the inspiring reflections on food-related matters elaborated in the dynamic field of food philosophy that the present contribution takes its point of departure. Relying on Ricœur's concern with the structure of lived experience, I suggest to consider the fruitful intersection between food philosophy and Ricœur's hermeneutical phenomenology. Rather than focusing on food as object of study, the attention will be oriented toward the discussion of the lived experience of being hungry, as a primary mode of being in the world, and of eating, as an act formed by inner physical and outer cultural and social mechanisms. Yet, given that the scope of Ricœur's philosophical thought is very broad, the present contribution does not seek to offer a comprehensive account of his work in the context of food philosophy. My claim is more modest. Following his work and taking a critical step beyond it, my article aims at considering Ricœur's heritage to innovatively improve

researches on food existentialism. His seminal phenomenology of embodiment and his attention to the elusive nature of time elaborated in his narrative hermeneutics, can be involved in the development of what I call an 'interpretive existential philosophy of being hungry and eating.' I will take a look at these moments by way of suggesting new directions for a critical thought on hunger and eating as existential problems which shape our perception of the world and the temporal dynamics of our personal and collective life. By applying Ricœur's oeuvre to the current debate in the philosophy of food, my essay will prove helpful to Ricœurian scholars interested in employing his thought to address prospects in philosophy and other fields that respond to emerging issues of importance. Yet, I believe that also philosophers of food can benefit from the extension of Ricœur's hermeneutic phenomenology to the field of food philosophy as long as it offers valuable methodological resources for exploring how food mediates our experience (see Kaplan 2020).

This article consists of two entangled sections. First, I will consider Ricœur's work *Freedom and Nature: the Voluntary and the Involuntary* (1966) as presenting relevant elements for opening up an inquiry of hunger and the eating as existential aspects defining the human condition. By focusing the attention on the corporeal involuntary, the phenomenological analysis of hunger and eating will be developed through the consideration of bodily needs, motives, and values. Hunger and eating will be analyzed as involved in the correlation between the involuntary bodily functions and the voluntary creative adaption to them. Then, I will shift the emphasis from the phenomenological approach to hunger and eating to the hermeneutical understanding of the connection between these dimensions and our experience of time. Developed in his monumental work *Time and Narrative*, Ricœur's conception of time as a combination of cosmological laws and lived experiences provides us with critical tools that help us think in a more consistent manner the differences and the continuity between clock-time and stomach temporality (see Boisvert 2006: 40), i.e., the quantitative and the qualitative component of the gustatory temporality.

1. *Hunger and eating at the edge of the will. From* Homo cogitans *to* Homo appetens

Ricœur has never elaborated a philosophy of food nor left any major work dedicated to nutritional issues. However, I believe that his thought can provide tools for exploring the meaningfulness of our food experiences. Arising in the midst of the problems of real life, i.e., in the conceptual confusion pool of everyday world and its multiplicity of conflicting meanings, Ricœur's oeuvre

offers us concepts and methodologies that help us to puzzle out the way we experience our bond with food as indicative of our relationship with our lived body and, by extension, with ourselves as acting and suffering beings living together with others in a socially shared environment (see Ricœur 1992: 23). Throughout his writings, Ricœur shows a constant interest in human agency. Denoting the basic manner in which human beings exist and inhabit the world, human agency is not understood as pure spontaneity; rather, human actions connect the voluntary and involuntary structures of our will, individual projects and worldly events. Actions involve "not only doing and making but also receiving and enduring, the latter being a joining of receiving and doing" (Dauenhauer 1998: 100). Following Ricœur's attention to the tensive structure of agency, I claim that Ricœur's thought can be significatively entangled in the development of a philosophical thought on hunger and eating as dimensions connected to the tension between passivity and activity characterizing our embodied existence. More precisely, Ricœur's early phenomenology of the will and his hermeneutical approach to the cosmic time and the phenomenological time of experience can find a renewed appropriation in the domain of food existentialism (see Kaplan 2019: 150). With that in mind, let us first focus on hunger and eating within Ricœur's phenomenological analysis of the corporeal involuntary.

Inserting his thought into the contemporary philosophical debates on the topic of the lived body, Ricœur's analysis of this subject finds its originality at the intersection of existentialism and phenomenology. Ricœur's aim is to develop an inquiry of human being's mode of incarnation in the world through a descriptive methodology. Otherwise put, according to Ricœur, the intuition of human being's incarnate existence inherited from existentialism has to be examined through a rigorous method that calls for a reconsideration of Edmund Husserl's descriptive phenomenology. Against the modern conception of the body as an obstacle to mind prefigured by the Cartesian Cogito and culminating in Husserl's *Cartesian Meditations* (1931), in his phenomenological analysis of human will Ricœur recomposes the integral experience of human being as marked by a relation of correlation (*corrélation diagnostique*) between the body and the intentional structures, bodiliness and consciousness. Contrary to any Cartesian dualism, which separates subjective mental thoughts from bodily movements, Ricœur stresses that there is "a single universe of discourse in which thought and movement would be homogeneous" (1966: 217). Thus, he claims that "the reconquest of the Cogito must be complete: we can only discover the body and the involuntary which it sustains in the context of the Cogito itself. The Cogito's experience, taken as a whole, includes 'I desire', 'I can', 'I intend', and, in a general way, my existence as a body" (1966: 9). Accord-

ing to Ricœur, human being is an embodied being whose life is characterized by the preservation of a fragile equilibrium between the subjective experience of the incarnate will and the objective aspects of the natural structure. Following his line of thought, the corporeal dimension is the primordial source of most original needs, motives, and organic values, arising from the spontaneous demands of life and echoing in the depths of our lived body. Let us examine now how these aspects of the corporeal involuntary relate to hunger and eating. The analysis of needs, motives, and values, enables us to understand the irrevocable intertwining between the carnal and the cognitive, the inner and outer horizons of our hunger and eating experiences.

(a) I argue that it is in Ricœur's discussion of the set of vital needs, that we can find elements for a first phenomenological description of hunger. Since we are born hungry and we have been hungry even before we can remember being alive or "gaining self-consciousness of our own pleasures" (Borghini 2017: 2), hunger can be coherently defined as our primary mode of being in the world. Phenomenologically speaking, hunger is not experienced from a third-person perspective. Rather, as all other bodily needs, hunger is felt within a first-person experience of the lived body in relation to the world. Hunger epitomizes, then, the relation between our bodily dependence on food and the social universe of which we are a part. Ricœur argues that hunger "does not reveal my body to me but through my body reveals that which is not here and which I lack. I do not sense contractions and secretions – I am aware of the I-body as a whole lacking" (1966: 91). Analogously to all other vital needs (thirst, sexual urges, etc.), hunger can be defined as a pre-representational state of outness. Through the consideration of "the adherence of affectivity to thought itself" (1966: 86), we can state that the felt need of hunger adheres to consciousness, opening it up to a description of intentional type. As Ricœur observes "to feel is still to think, though feeling no longer represents objectivity, but rather reveals existence" (1966: 86). Hunger is, then, an indigence and an exercise directed toward something, "it is a consciousness of…an impetus towards…Even without an image of bread, my hunger would still carry me beyond myself" (1966: 90). Otherwise put, hunger carries an affective-intentional load. By rejecting the model of felt needs employed by objectivism, Ricœur invites us to think hunger neither as an inner sensation nor as part of a stimulus-response pattern, but as a referential or transcending behavior linked to conscious acts intentionally directed towards the world. As such, expressing the need to eat, hunger makes us experience our interdependence with the world, exposing us to the otherness of what is edible. This exposition might lead us to a pleasant fulfillment, but also to the risk of being disgusted or even to the danger of being poisoned. In this sense, hunger reflects human being's deepest vulnerability. It is in this context that Ricœur

concludes: "the will completes the separation of experience and need: while the impetus can be mastered by the will, the lack always remains un-coercible – I can refrain from eating, but I cannot help being hungry" (1966: 91). As such, "hunger is more fundamental than pleasure for the human condition" (Borghini 2017: 3). Therefore, the satisfaction of hunger "is more fundamental than sex. In the life of the individual organism it is the more primary and recurrent want, while in the wider sphere of human society it determines […] the nature of social groupings, and the form of their activities take" (Mintz 1985: 4).

(b) Our vital needs are bound to bodily motives. Indeed, our body is not just the source of the most original needs, but also of primordial motivations deriving from the spontaneous demands of life. Considered in the context of our corporeality, motivation is an inner move emerging from the deepest realm of our lived body. More precisely, according to Ricœur, motivation is an intentional stream that inclines the will to decide for something "in order to" as well as "because of". Thus, every motive is a motive for a decision that inclined the will towards its projects. As Ricœur stresses, "the circular relation of motive to project demands that I recognize my body as body-for-my-willing, and my willing as project-based-(in part)-on my body" (1966: 85). Considered as an inner urge, hunger is the main motivation to seek food and eat. In order to understand this point, we have to explain how hunger can give us a reason on the basis of which we are motivated to eat. The fact that feeling hunger is inherently motivating means that hunger is not a mere sensation, but it is an affective reaction because it involves "changes in affect, that is, positive or negative hedonic feelings and relatedly, felt states of attraction or aversion" (Ombrato, Philips 2021: 520). The agent's motivation to eat is not just instrumental as directed to the alleviation of the unpleasant sensations associated to hunger at the personal, experiential level. Rather, it also involves a positive attraction we feel towards food, that is, an anticipated pleasure to satisfy our hunger. Thus, "hunger facilitates the elicitation of appetite, a felt desire for food or attraction to the prospect of eating, and that such phenomenon is recruited by hunger to further its motivational role, so that, ordinarily, we are at once driven by hunger and drawn by appetite" (Ombrato, Philips 2021: 518). Hunger and the related act of eating encompass the physiological, psychological, hedonistic, and broadly socio-cultural aspect of our existence. In this sense, hunger as a motive is interwoven with countless other motives such as those concerning the meal type, the type of food, the quality and the quantity of food, etc. Therefore, besides the biological mechanism, certain other personal, social, cultural, and psychological factors are connected to hunger. For example, given that I am vegetarian, my hunger is also to my ethically motivated refusal to consume meat and to hurt animals. Hunger is, then, involved in a complex framework of

other motives and it "would be futile to try to enumerate every food-motive for every food action" (Kaplan 2019: 153). In hunger and eating we find a "point of convergence between the rational and emotional aspects of human agency, as well as the point of convergence between humans as biological organisms and as expressions of culture" (Borgini 2017: 3). As such, motives relate to food choices and eating habits as "immensely important adornments on an inescapable necessity" (Mintz 1985: 3).

(c) The body is also the source of the most original living values. Following Ricœur's lead, the body reveals the primordial layer of values: organic values. Although these are all directed toward the organic well-being, they are heterogenous and concern for example assimilation, security, exercise, and rest. A value can reach and incline me only if it impresses my sensibility as a dignifying motive. Thus, Ricœur explains that since the body is the basic involuntary and the basic fundamental source of motivation, all other values are elaborated in relation to it (see 1966: 122). As he puts it, as long as the body is the affective medium of all values, these "assume a serious, dramatic significance through comparison with the values that enter history through my body" (1966: 85). In this sense, organic values open up the space for the level of history, i.e., for the meaningfulness of the cultural and social values. Ricœur invites us to think that even though we can actualize organic values in different ways, we need to attend all of them in some balanced fashion as necessary conditions of our existence. In the case of hunger, it inclines our will to pursue something perceived as good to eat for us. For example, hunger makes us perceive bread as good with its agreeable taste and its suitability for assuaging my lack of food. Therefore, food-values are linked to the so-called "omnivore's dilemma" (see Fischler 1988; Pollan 2006). Since human beings are omnivore beings who have to eat in order to survive, in virtually any circumstance in which foods are supplied they are confronted with the choice of whether or not to eat them, and which ones to consume according to their personal and cultural values. Food-values are involved into the environment where we live. More precisely, we can state that our food-values are included within a complex gustatory environments in which they are susceptible to external influences and internal forces that are not completely under our control. Otherwise put, food-values are linked to preferences that consciously or unconsciously influence our daily food choices such as "price, convenience, taste, health, appearance, familiarity, novelty, mood, diet, and ethics" (Kaplan 2019: 154). These values can change through time depending on the evolution of the collective culinary imaginary and can also be manipulated or nudged by the society.

Ricœur's phenomenological description of the corporeal involuntary, that is, of the lived body as primordial source of needs, motives, and values, leads us

to consider hunger and the consequent intake of food through the act of eating in the context of the dialectic between involuntary and voluntary dimensions of human will. Following Merleau-Ponty's refusal of Sartre's voluntaristic excess (see Marsh 2008), Ricœur replaces the existentialist dualism of facticity and transcendence with a two-dimensional unity of human will in which the involuntary and the voluntary are complementary. However, unlike Merleau-Ponty, Ricœur claims that the description of human being's embodied existence cannot be separated from its ethical and theological implications. With reference to Marcel's Christian existentialism, Ricœur conceives the lived body as part of human freedom by defining it in terms of an "incarnate freedom" (1966: 33). In the context of our relationship with food, we can state that on the one hand, human beings are inescapably bound by hunger and eating as involuntary dimensions of the existence. On the other hand, though, our fundamental voluntary projects inform how we live hunger and the way we look for food. Let me explain this point further. Certainly, hunger and eating are the primal marks of life evoking the world as a great cosmic banquet. From a biological perspective, human beings have to eat as well as all other living beings in order to survive. Nevertheless, for human beings hunger and eating cannot be limited to the instinctive-animal level. These existential dimensions reflect human beings' original relationship with their own personal experiences, with the environment, with other living beings and inanimate things. Human beings are social beings whose behaviors are shaped and apprehended within the socio-cultural context. Hunger and eating figure, then, into meaningful agency taking on different meanings related to human beings' own understanding of their gustatory experiences. As creatures that seek meaning, produce meanings, and yearn from meanings, human beings have the power to make meaningful what in Heideggerian terms can be defined as the "worldhood of food" (see Kaplan 2019: 169). However, human beings are not transcendental masters of meanings since they are marked by their constitutive bodily finitude and existential frailty. Given that our will is embodied, the act of willing consists at once in the realization of freedom and in the reception of necessity. Human beings are free agents, that is, they are above the necessity of physics since they can act outside the pure scope of laws of mechanics. But, at the same time, they are determined by bodily necessities and by circumstances which are beyond total control. By applying these reflections to the understanding of hunger and eating, we can state that we necessarily need to eat in order to survive, but we are not forced to eat a determined food, like koalas holding into branches of an eucalyptus three. Yet, even if we can freely choose our foods, we are not disengaged from our body and external influences. Thus, our freedom is bounded freedom governed and limited by the body and by the context of its

occurrence. Neither freedom can be content with simple acquiescence to the necessity of nature, nor it corresponds to pure arbitrariness. For example, my freedom to choose what to eat does not mean that I am free to eat all candies of the Halloween party. Certainly, I can choose to do that if I do not matter for my health. Contrary to Sartre's theory of absolute freedom and his thesis according to which human beings are "condemned to be free" (1956: 567), Ricœur conceives freedom as a 'bound freedom' inseparable from the opposition and the struggle with the involuntary, rejecting all negative sense of it as a state of condemnation. In this perspective, Ricœur displaces the question of the nature of human being from the quest of epistemological certainty to that of life as enjoyed: "I sense myself alive before I know myself as an animal" (1966: 411). Human existence is more a matter of what I feel and what I can do, rather than "I think". Interpreting the dialectic between hunger and eating in this context, hunger is more original than the act of eating and thinking. In conclusion, I claim that Ricœur's phenomenology of the embodied will allows us to coherently affirm 'I am hungry, therefore I am.' Ricœur's phenomenology opposes to *homo cogitans*, whose emblem is the mind, *homo appetens* existing as a betweenness of affection and intention.

2. *The temporal mediation between clock and* gaster*: the gustatory time*

We have just seen that Ricœur's phenomenology of the will presents significant elements to think hunger and eating as existential dimensions intertwined with our lived body. Having analyzed the phenomenological insights, we should turn our attention now to the productive tension between hunger, eating, and the experience of time. In other words, following Ricœur's line of thought, we have to move from a descriptive inspired phenomenology of hunger and eating to a hermeneutical phenomenology of the experience of these existential phenomena as linked to the temporal dimension of human existence. The phenomenological approach to hunger and eating becomes, then, more complex and nuanced. It is my contention that the movement from descriptive phenomenology to hermeneutic phenomenology shows a methodological and epistemological tension that inwardly affects the philosophical approach to hunger and eating. The changes of method implied by Ricœur's evolution from a descriptive phenomenology to an explicitly hermeneutic one, allow us to move from the conception of hunger and eating as existential problems connected to bodily necessities, in which a negative sense of constraint prevails, to the outline of hunger and eating as linked to the productive power of human experience, in which a more positive conception of these dimensions is at stake. Yet, insofar as Ricœur erects hermeneutics on the basis of phenom-

enology, we can stress that the two different methodological approaches to hunger and eating are dialectically related.

The phenomenological description of human being's nature as determined by a discordant-concordance between the poles of the voluntary and the involuntary, freedom and necessity, is recast in Ricœur's monumental three volumes of *Time and Narrative*. His study of the relationship between time and narrative offers a provocative framework for examining the connection between clock-time and stomach-time (see Boisvert 2006: 40), that is, between what one can consider as the quantitative and qualitative temporal dimensions of our experience of being hungry and eating. In the first volume of *Time and Narrative* (1984), Ricœur focuses his attention on the antinomies affecting the conception of time. He takes as his starting point two of the most influential philosophical reflections on time: Augustine's question about how to measure time in Book XI of *Confessions* and Aristotle's study of emplotment (μύθος) in *Poetics*. Let me briefly review Ricœur's argument as essential for understanding the intertwining between clock and stomach time in the context of food existentialism. Ricœur begins his analysis of time with the reconsideration of Augustine's famously statement: "what, then, is time? (*Quid est enim tempus*) I know well enough what it is, provided that nobody asks me; but if I am asked what it is and try to explain, I am baffled" (11, 14:17). Augustine's question is to understand how time can be said to be since the past is no longer, the future is always not yet, and the present is always not always. Augustine's solution is to suggest that the temporal experience is a threefold present grounded on the distention of the soul across time. As he puts it, "some such different times do exist in the mind, but nowhere else that I see. The present of past things is the memory; the present of present things is direct perception; and the present of future things is expectation" (11:20). Thus, the distention (*distentio*) of the soul is stretched by the separate intentions (*intentio*) of expectation, attention, and memory. According to Ricœur, Augustine's threefold present does not resolve the enigma of time by displacing it to an internal problem. In order to think the relationship between the time of the soul and the time of the cosmos, Ricœur turns to Aristotle's *Poetics*. Whereas in Augustine's meditations on time the reign of discordance dominates over concordance, Aristotle's theory of emplotment (μύθος) offers a structure for thinking the concordance of events in an unfolding narrative over discordance. The emplotment is understood as an act of configuration in which the unity of the plot across time and the events, that make up the various components of the action, are balanced together. Therefore, emplotment is "the synthesis of the heterogeneous" (Ricœur 1984: 66). The aporia of time experienced as a concordant discordance in Augustine's thought, finds its remedial counterpart in Aristotle's idea of narrative

mimesis as discordant concordance. Aristotle's theory provides the guiding principle for the entirety of Ricœur's trilogy of *Time and Narrative.* The critical analysis of Augustine's philosophical reflection on time and Aristotle's discussion of poetics lead Ricœur to affirm that "time becomes human to the extent that it is articulated through a narrative mode, and narrative attains its full meaning when it becomes a condition of temporal existence" (1984: 52). Going further than Ricœur, I claim that narrative is not only the medium through which we understand ourselves as living, acting, and thinking beings. Rather, narrative can be also applied to the understanding of being hungry and eating as rhythmic temporal experiences expressing our relation with food.

Hunger and eating are involved in the external structures of clock time, expressed by devices such as calendars, that split time up into homogeneous units such as days, months, years. As Ricœur puts it, the calendar "cosmologises lived time and humanizes cosmic time. This is how it contributes to reinscribing the time of narrative into the time of the world" (1988: 109). Through the clock time, the time of all human actions, whether individual or social in scope, is inserted into chronological and measurable time. Thus, clock time enables us to understand time with measurement. What the clock time measures is a quantity of duration, that is, a stretch of time. Clock time does not measure the endurance of time itself. Rather, it measures the motion of the internal components of the clock in terms of conventional intervals. More simply, the clock measures motions of temporal segments of varying lengths (e.g., seconds, minutes, hours). Therefore, our understanding of clock time tells us the measurement of intervals which have been commonly agreed as representing the stretches of time. Clock time stands in a close relationship to hunger and eating. The quantitative measurement of time provides the rhythm to ritualized celebrations, which differ from one culture to another and that are associated with community's ritual meals. For example, in the United States the last Thursday of November is the day of Public Thanksgiving, which is celebrated with a traditional meal of turkey, cranberry jelly and pumpkin pie, whereas in Japan each year on March 3rd is the Hinamatsuri (雛祭り) or Girls' Day celebrated with typical Japanese food. Clock time is, then, connected to our collective gustatory identity as part of a living culinary tradition. Yet, clock time is tied up to our daily eating routines, according to our personal and collective life schedules. Specifically, in our society the time of the clock triggers our breakfast time, lunchtime, and dinnertime. However, the time-based decision to eat does not always correspond to the physiological need to eat, i.e., to the hunger drive. Although the clock time regulates our daily schedules, it does not imply a direct correspondence between the expressions "it's time to eat" and "I am hungry enough to eat now" (Somov 2008: 24). There are also considerable differences among individu-

als concerning the relation between eating and clock time: some people miss breakfast, while other have their main meal at noon rather than in the evening. These differences are determined by personal preferences, schedules, as well as cultural patterns. Linked to the clock time, hunger and eating appears as daily repetitions of the same habits. Contrary to the modern divorce between time and space announced by Newton's pronouncement that "time exists in and of itself and flows equably without reference to anything external" (Greene 2004: 46), clock time makes time understandable for us since it measures the endurance of something in space. In his theory of narrative, Ricœur highlights the connection between temporality and context, time and space, in describing the qualities of a coherent and meaningful narrative by means of mimesis. Indeed, according to Ricœur, time cannot appear separately from the contents of experience (see 1988: 23). With reference to hunger and eating, clock time is connected to the place where we live our hunger drive and perform the act of eating. The timing of our intake of food serves a social function for the formation of our personal and collective identity. Regulated by the clock time, hunger and eating are intertwined with the larger time frame of nature, of seasons and of the alteration of day and night, but also of the cultural and historical contexts. In other words, we can state that clock time measures hunger and eating including them in the framework of cultural and social phenomena. It does not mean that we always eat with someone. Even when we do not share food, as an act regulated by clock time corresponding to shared conventions, eating is involved in the social sphere. Therefore, the time of the clock is inherently connected to the rhythms associated with intersubjectivity, culture, traditions, as well as to the register of language and of the symbolic. Clock time is, then, an essential component of the gustatory time in which our gustatory identities are formed through events, stories, and common habits. Following Ricœur's line of thought, we can conclude that a reflection on the clock time with reference to hunger and eating does not involve just a question of "who I am in time", but also "who I am *with* through time", that is, it implies the problem of gustatory identity as temporally and socially constituted by external processes.

Hunger and eating are not regulated just by the quantitative time of the clock as external temporality. Through a coordinated biological system, our body performs certain tasks through what is called the bodily microbiological clock. Given that our body relies on the natural world, bodily time is adapted to manage daily environmental changes, such as the atmospheric lightness and darkness caused by the cycle of Earth rotation. There is, then, a complex synergy between the rhythm of our body and the cosmological flow of time, between the qualitative time of the body and the quantitative time of the clock. Among the bodily rhythms, stomach time deserves a special attention as an embodied time

asking for a necessary encounter with the external world. From the stomach time a centrifugal and a centripetal back and forth movement towards the otherness of the world begins. Contrary to other corporeal rhythms such as the circadian rhythm or the leaver metabolic rhythm, stomach time is personally lived. Indeed, the stomach time is a felt experience arising from the incessant demand for something lacking. The stomach time and its fulfillment testify human being's primordial relationship with the world. More precisely, the stomach time makes us experience an "immersion of our sensibility [...] and the agreement that we expect to find at first glance between our needs and the world" (Pelluchon 2019: 33). Therefore, through the stomach time we experience at once our dependence with respect to the world and the possibilities that this world, in which we move and constitute ourselves as social beings, gives to us. The satisfaction of hunger goes beyond the need to ingest some food in order to survive and not to perish. However, our relation of dependence and interaction with the world, might also lead our stomach time to be a context of pain and suffering. In short, the stomach can lead both to experience joy and pain, pleasure and sorrow. In virtue of our lived sensibility, the stomach time uncovers the existential character of being hungry and eating in relation with our being-in-the-world as embodied beings involved and touched by exterior things. For example, our stomach time becomes a time of pleasure when we come into contact with some foods that satisfy hunger. But it can be also a time of pain, if we do not come into contact with desired types of food or even with food itself in extreme cases. Yet, the stomach time can be experienced as an empty-fulness or as an ongoing emptiness as it is in the case of some eating disease such as anorexia, bulimia, and obesity. As such, the qualitative time of the stomach manifests through fullness or emptiness, slowness or rapidity, satisfaction or lack. Stomach time can be transformed also into a pure egoistical time accompanied by the popular statement that 'a famish stomach has no ears'. This sentence does not just have a literal signification in the moral sense that if we are hungry, we are unable to listen attentively and we might act in an egoistic way. Beyond its empirical meaning, this expression highlights the fact that the potentials of human conscious life depend on the meeting between the basic bodily needs and the external world. The stomach time reminds us, then, of our material dependence and permanent vulnerability.

Certainly, the stomach is an organ that has been often ignored by philosophers. The importance of the stomach as a current topic in food philosophy is grounded on the reconsideration of the body as one of the most important themes of contemporary philosophy. Specifically, we can develop a philosophical approach to stomach time only after what Ricœur calls the total reconquest of the Cogito (1966: 9), that is, after all rejection of the Cartesian dualism and the modern conception of human being as a thinking mind characterized by

the status of self-sufficiency. Coherently with the critique of modern philosophy as "mind-intoxicated" (Boisvert 2006: 42), I believe that Ricœur can be considered as a stomach-friendly thinker. The fact that we are stomach-endowed beings is not something negative or as a disgrace to be hidden in favor of our thinking mind. Put differently, our mind and consciousness only last as long as the stomach allows (see Minister 2015: 30). In this renewed perspective, the stomach time is not a burden, but it opens up our first intentional and interactive encounter with the world, making us understand temporality as linked to "opportunities surging forward" (Boisvert 2006: 42). As social beings, we are temporally intertwined with the gustatory environments in which we are embedded. Our stomach time is involved in the conjunction of actualities and possibilities evolving in response to contexts and in light of influences. Thus, the construction of the gustatory time deals not exclusively with the preservation of our identity in the process of time, i.e., in the chaining of temporal events regulated by clock time, but also with the productive power linked to a human being's stomach-situatedness into a shared world. As Ricœur suggests, since my life story is caught up in the stories of others, the gustatory time of each human being cannot be just individual. Consequently, as well as the clock time, stomach time as part of our identities cannot be taken separately from the social nexus in which human beings exist. Stomach time is, then, inseparable from intersubjective practices and the active critical appropriation of the gustatory environment in which we live. We do not simply inherit a gustatory tradition. As interpreting beings, we can take a different attitude on what, when and how we eat, that is, we have a productive capacity to reflect and imagine new gustatory possibilities over time (see Borghini, Piras 2021).

3. *Conclusion*

In this article I have investigated the opportunity of a fruitful development of Ricœur's thought in the direction of food philosophy. In doing so, I have proposed a critical reading of his hermeneutical phenomenology and shown how it can contribute to the formulation of a philosophical reflection attentive to hunger and eating as existential dimensions of human life. Ricœur's work offers insights of hunger and eating that enable us to develop an interpretative existential philosophy of these key aspects of our being-in-the world as relational hungry and eating beings.

My interest was primarily focused on Ricœur's phenomenology of embodiment as presenting relevant elements for a descriptive analysis of the hunger drive and the act of eating. I have drawn attention to these aspects as involved in the lived experience of human being as an incarnate cogito, that is, as a rela-

tional subject whose embodied experience is always and already in touch with the otherness of the world. By focusing on Ricœur's examination of the corporeal involuntary, I explained how hunger and eating relate to bodily needs, motives, and values, stemming from the spontaneous demands of life. In this context, hunger and eating arise as structured within the "double allegiance" of the human embodied condition (Ricœur 1992: 111): on the one hand hunger and eating are bound to the laws of the natural world, and on the other hand, these existential dimensions can be mastered through our limited freedom. Hunger and eating are, then, understood as structured through a series of dialectically related dualities, such as passivity and activity, identity and diversity. Ricœur's phenomenology of embodiment can be reinterpreted in the development of an existential philosophy of food, aiming at presenting hunger and eating as interrelated moments of our existence characterized as a mixture between the involuntary bodily functions and the voluntary adaptation to them.

The phenomenological-existential analysis of hunger and eating has been further developed through the correlated interpretative reflection of these dimensions as experienced over time. Through the reconsideration of Ricœur's theory of narrative, I explored hunger and eating as involved within the dialectic between clock-time and stomach-time. Contrary to the modern perspective of human being as un-hungry mind, the reconquest of human being's integral experience forms the basis for the development of any philosophical approach to hunger and eating as existential dimensions embedded in our personal and collective experience of time. The possibility of a dialectical intertwining between clock time and stomach time, interpreted as the quantitative and the qualitative dimensions our experience of being hungry and eating, has led to the formulation of what I have called the gustatory time as a temporality situated in the shared worldhood of food.

Hunger and eating involve a back-and-forth movement from ourselves towards the world, i.e., a give-and-take between the flourishing of our lives and the circumstances within which our aspirations can be realized. Human being as a stomach-endowed being, that is, as *homo appetens*, is an effort and a desire to exist. This dynamic potency that expresses itself through hunger and eating, opens up to our existence as a creative recipe in which we project ourselves through the ongoing tension between circumstances and new gustatory imaginaries.

Maria Cristina Clorinda Vendra
Institute of Philosophy of the Czech Academy of Sciences, Prague / Institute of Philosophy of the Slovak Academy of Sciences, Bratislava
vendra@flu.cas.cz

References

Boisvert, Raymond, and Heldke, Lisa, 2016, *Philosophers at Table. On Food and Being Human*, Reaktion Books, London.

Boisvert, Raymond, 2006, "Clock-Time/Stomach-Time", in *Gastronomica*, 6, 2: 40-46.

Borghini, Andrea, and Piras, Nicola, 2021, "On Interpreting Something as Food", in *Food Ethics*, 6, 1 (online first, December 2020),

Borghini, Andrea, 2017, "Hunger", in D. Kapla and P. Thompson, eds., *Encyclopedia of Food and Agricultural Ethics* (2nd ed.), Springer, New York: 1-9.

Dauenhauer, Bernhard, 1998, *Paul Ricœur: the Promise and the Risk of Politics*, Rowman and Littlefield, Lanham.

Kaplan, David, 2019, *Food Philosophy: An Introduction*, Columbia University Press, New York.

Marsh, James, 2008, "Ricœur's Phenomenology of Freedom as an Answer to Sartre", in Kaplan, David, ed., *Reading Ricœur*, State of New York University Press, New York: 13-29.

Minister, Stephen, 2015, "Food, Hunger and Property", in J. Dieterle, ed., J*ust Food: Philosophy, Justice and Food*, Rowman & Littlefield, London: 21-38.

Mintz, Sidney, 1985, *Sweetness and Power. The Place of Sugar in Modern History*, Penguin Books, New York.

Ombrato, Michele, and Philips, Edgar, 2021, " The Mind of the Hungry Agent: Hunger, Affect and Appetite", in *Topoi*, 40, 3: 517-526.

Pelluchon, Corine, 2019, *Nourishment. A Philosophy of the Political Body*, Bloomsbury, London.

Pollan, Michael, 2006, *The Omnivore's Dilemma. A Natural History of Four Meals*, Penguin, New York.

Ricœur, Paul, 1966, *Freedom and Nature: the Voluntary and the Involuntary*, Northwestern University Press, Evanston.

Ricœur, Paul, 1967, *The Symbolism of Evil*, Harper and Row, New York.

Ricœur, Paul, 1984, *Time and Narrative. Vol I.*, University of Chicago Press, Chicago.

Ricœur, Paul, 1988, *Time and Narrative. Vol. III*, Chicago University Press, Chicago.

Ricœur, Paul, 1992, *Oneself as Another*, University of Chicago Press, Chicago.

Sartre, Jean-Paul, 1956, *Being and Nothingness*, Washington Square Press, New York.

Somov, Pavel, 2008, *Eating the Moment*, New Harbinger, Oakland.

Vendra, Maria Cristina, 2020, "Interpreting the Natural Environment. Paul Ricoeur's Directions for an Eco-Hermeneutic Phenomenology", in *Discipline Filosofiche*, 30, 2: 261-279.

Focus

The conceptual framework of persons. A metaphilosophical investigation

Introduction

Danilo Manca, Giacomo Turbanti

At the end of his essay *Philosophy and the Scientific Image of Man*, Wilfrid Sellars famously insists on the need to integrate the categorial framework of contemporary sciences with the conceptual framework of persons. The former is characterized by the languages of scientific theories, whose objects are theoretical and whose explanations are postulational. The latter is characterized by the language of community and individual intentions, whose objects are manifest and whose explanations are normative.

Since Sellars' original diagnosis, the clash of the two conceptual frameworks has grown more and more dramatic. On the one hand, the very legitimacy of the manifest image of common sense through which individuals share their intentions and project their actions is threatened by the reductivist claim of objectivity of the scientific image. On the other hand, by obliterating the normativity of the framework of persons, the scientific image risks offering a representation of the world devoid of the reasons and justifications that could be incorporated into our everyday life.

This Focus proposes looking at this problem from a metaphilosophical point of view, and embarking on an investigation of the different metaphysics, epistemologies and logics that characterize the two different conceptual frameworks. Such an investigation will hopefully unveil the peculiarity of the conceptual framework of persons for other views of the world and illuminate to what extent the acknowledgement of this peculiarity contributes to the reflection the philosophical inquiry dedicates to its specific practice and conditions.

The word "metaphilosophy" is often criticized, even by those who launched it. Timothy Williamson (2007: IX) rejected it because "metaphilosophy sounds as though it might try to look down on philosophy from above or beyond", while the philosophical reflection on the activity of philosophizing is "automatically part of philosophy". Richard Rorty (1992: 374) already pointed out that "questions about the method of philosophy", or "the nature

philinq X, 1-2022, pp. 43-48
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.351

of philosophical problems" are "likely to prove unprofitable". Equally skeptical, Bernard Williams (2006: 169) radically excluded that philosophy is "at its most interesting when it is talking about itself". More generally, going back to the original meaning of the term "philosophy", one could legitimately assert that philosophy is always accompanied by reflection on itself precisely insofar as it is intent on seeking truth, wisdom, virtue. And yet, as McGinn (2002: 199) emphasized, the inquiry into the nature of philosophy is "perhaps the most undeveloped part of philosophy".

By explicitly adopting a metaphilosophical approach to the investigation of the conceptual frameworks of persons, we aim to show that the act of positing itself as the subject of investigation is not self-referential, not a moment in which philosophy suspends any investigation of the world; on the contrary, we intend to show that any investigation philosophy develops, is also, inevitably a stance on what philosophy is. Metaphilosophy is therefore an approach that pays attention to a practice that is always at stake in the act of philosophizing, but often remains unexpressed: that of defining oneself to be able to tackle the problems that arise from time to time.

The metaphilosophical investigation on the conceptual framework of the person can be addressed from different points of view and different philosophical traditions. This collection alone features authors and contributions variously linked to pragmatism, phenomenology, the analytic philosophy and critical theory. Quite interestingly, this sort of investigation favors the merging of diverse approaches and the development of more comprehensive and conscious perspectives.

In the first contribution, Giacomo Turbanti describes the semantic implications of the clash between the images as a metaphilosophical problem. He argues that while the clash has often been discussed with regard to its ontological impingements, it should be seen instead as generated by the incompatibility between the two different collections of categories that articulate the conceptual frameworks of what Sellars called the manifest image and the scientific image. As a consequence, he suggests, the clash can be understood as raising two important questions. The first one is about what categories should be adopted for representing the world. The second one is whether philosophy is entitled to those categories or should better give up on the task of representing the world entirely. The latter question is particularly problematic in the context of Sellars' characterization of the "perennial" tradition in philosophy as endorsing the manifest image as real and the concept of a person as the foundation of the normativity of the space of reasons.

Sellars however also thought that the concept of a person as a subject of ought-to-do-rules could be integrated in the scientific image by undergoing a

semantic process of recategorization. David Landy's paper is devoted to the analysis of how Sellars developed his notion of person by confronting Kant and the problems of the Transcendental Deduction. Landy draws a parallel between the Kantian recognition of the analytic unity of apperception and the Sellarsian view that rule-governed practical reasoning depends on the unity of the subject of such reasoning. However, Kant also maintained that the analytic unity of apperception is possible only under the presupposition of a certain synthetic unity, which in turn requires the definition of a collection of categories as the fundamental rules of synthesis. While Landy argues that Sellars ultimately follows Kant also in accepting the synthetic unity of apperception as the condition of any representation of an object, the question remains whether the categories of the scientific image are suitable to provide the rules for the appropriate synthesis.

In his paper, Carl Sachs explores the extent in which behavioristic psychology and cybernetics could have inspired Sellars in envisioning a path for the recategorization of the concept of a person in the scientific image. Sachs describes Sellars in his early work as searching for a non-psychologistic (or, better, a psychologically correct) way to naturalize Kant's idea of epistemology as a synthetic *a priori* enterprise. His methodological adoption of pragmatism and behaviorism should be read in alignment with this strategy. The mature conception of this strategy hinges on the notion of "picturing", which is a key element of the Sellarsian account of intentionality in the scientific image. Sachs argues that Sellars would have developed this notion by reflecting on the cybernetic theory. According to what he calls "cybernetic behaviorism", how a cognitive system produces representations of the world is determined and cannot be understood independently of the complex material dynamic of feedback interactions between the system and its environment. Sachs suggests that a naturalization of normativity could be pursued by integrating this cybernetic analysis of intentionality with an account of how multiple systems could triangulate their behavior and coordinate with each other. According to this account a person could be recategorized in the scientific image as a cybernetic system that can reciprocally triangulate its behavior with other such systems.

The determination of the appropriate collection of categories with which representations of the world are produced in the scientific image is not the only subject that generates metaphilosophical problems about the conceptual framework of persons. This is obviously because persons are not simply representational systems. The next couple of essays discuss the emotional dimension of persons and the problem of investigating affectivity from a metaphilosophical perspective.

Peter Olen draws an interesting parallel between Sellars' mostly rationalistic characterization of human behavior as being governed by conceptual norms in the space of reasons and de Laguna's alternative conception that, while sharing relevant thematic and historical connections with the Sellarsian enterprise, reevaluates the roles of emotions and affects in the definition of persons. Olen focuses on moral actions and notices that Sellars ultimately provides an account of morality in which emotions are only considered as states that contribute to a causal explanation of agency. What is interesting of de Laguna's approach, in Olen's view, is that he managed to ground an account of the emotional dimension in the same behavioristic psychology that Sellars exploited instead as a methodology for a more comprehensive account of the intentionality of inner states. Behaviorism allowed him to work with the concept of a group mind in a naturalistic framework and describe the impact of emotions on those felt obligations that shape our social cognition.

Íngrid Vendrell Ferran explores Max Scheler's metaphilosophical view, by focusing on his thesis according to which philosophical knowledge presupposes a moral attitude. Scheler sees the philosophical attitude as determined by an act of upsurge that invests the entire personality of who she wants to be, or means herself as, a philosopher. At first, Vendrell Ferran focuses on Scheler's conviction that focusing on the type of person a philosopher will allow us to find out the nature of the object of philosophy itself. After emphasizing the sharp difference that, by contrast to Husserl, Scheler considers to exist between sciences (rigorously in the plural) and philosophy, understood as the capacity of intuiting essences, Vendrell Ferran deals with love, self-humbling, and self-mastery as the moral preconditions of philosophical knowledge. She criticized Scheler's essentialism insofar as it presupposes the adhesion to controversial metaphysical claims and proposed a reinterpretation of this affective categories in the terms of the debate on virtue epistemology.

The last two articles tackles the metaphilosophical issues underlying the essay *Philosophy and Scientific Image of Man*, i.e. the idea that "the aim of philosophy, abstractly formulated, is to understand how things in the broadest possible sense of the term hand together in the broadest possible sense of the term" (Sellars 1963: 1). The attempt at elaborating a stereoscopic view of both the conceptual framework of persons and that of sciences, without reconciling the manifest and the scientific image of man, is rooted in this thesis.

Danilo Manca takes part in the thriving debate on the match and mismatch between Sellars and Husserl with a metaphilosophical aim: to assess whether the Husserlian notion of "life-world" could be helpful for a philosophical theory that assigns a primacy to the scientific view of the world (as

Sellars did) when it comes to establish what exists, and accordingly, what "reality means". After excluding that the elaboration of a theory of the life-world necessarily entails the endorsement of the primacy of the manifest over the scientific image of the human being in the world, Manca introduces the standard Copenhagen version of the quantum physics to defend a pragmatic conception of realism. This allows Manca to contest two assumptions made by Sellars: the first is that reality cannot be conceived as stratified, the second is that the term "phenomenon" has to be understood exclusively in the supposedly Kantian sense of "illusory appearance". Danilo Manca shows that an abandonment of these two assumptions by a postulational attitude brings Sellars' and Husserl's perspective closer together, provided that we challenge Husserl's conviction that the "technization" of scientific inquiry entails a philosophical regress of the image of nature.

In the last essay of this Focus Paul Giladi argues that in his attempt at integrating stereoscopically in one unified and coherent image the conceptual framework of persons with a "Perceish" discourse, which construes everything in a purely naturalistic descriptive terms, what Sellars has carried out is to adopt a "*negative* dialectical resolution" of the clash between the manifest and the scientific images of man in the world. Following O'Shea and Christias, Giladi holds that Sellars' stereoscopic vision is construed as a functionalist naturalism integrated by a normative approach, insofar as persons are seen as logically irreducible but causally reducible to the descriptive categories of science. After reconstructing the peculiarity of Adorno's conception of dialectics in his reversal of Hegel, and hypothesising Adorno's criticism of Sellars's physicalist ontology, Giladi tries to envisage a left-wing Sellarsian response, by emphasizing that, in a curiously Hegelian fashion, Sellars explains that when he uses the analogy of the stereoscopic vision, he sees the manifest image as not overwhelmed in the synthesis. To Giladi, this means that Sellars's *Aufhebung* of the tension between the manifest and the scientific image points to a polychromatic, republican pluralism, rather than a monochromatic, imperialist monism. In other words, as deVries suggests, the relation between the two frameworks is a matter of mutual accommodation, not a mere dominance of one over the other. And yet, by so doing, according to Giladi, Sellars somehow, surreptitiously, overcome its fear of non-identical thinking, and leaves that the conceptual framework of persons corrosively and latently works on the tendency of the scientific image to assert its primacy over the manifest.

Danilo Manca, Giacomo Turbanti
University of Pisa

References

McGinn, Colin, 2002, *The Making of a Philosopher*, Harper-Collins, New York.

Rorty, Richard, 1992, "Twenty-five Years After" in Rorty, ed., The Linguistic Turn: Essays in Philosophical Method, University of Chicago Press, Chicago, 371-4

Sellars, Wilfrid, 1963, "Philosophy and the Scientific Image of Man" in Science, Perception and Reality, Ridgeview Publishing Company, Atascadero, 1-40.

Williams, Bernard, 2006, *Philosophy as a Humanistic Discipline*, Princeton University Press, Princeton.

Williamson, Timothy, 2007, *The Philosophy of Philosophy*, Blackwell, Oxford.

Some remarks on the categories of the manifest image

Giacomo Turbanti

Abstract: This paper addresses the question whether or not philosophical discourse can avail the categories of the scientific image. I argue that the clash of the images is better understood on the semantic rather than the ontologic level and that it results from the challenge to the representational adequacy of the categories tha articulate the conceptual repertoires of the manifest image. A challenge that will be met by a succesful recategorization of the concept of a person in the scientific image. I suggest some reasons to believe that such a recategoritazion is possible in principle without dismantling the philosophical discourse.

Keywords: Metaphilosophical Problems; Manifest and Scientific Image; Powers; Categories; Normativity.

1. *Introduction*

Bluntly put, the problem that Sellars described as the clash between the manifest image and the scientific image is the problem of whether or not philosophy is entitled to deal with the human being by its own means. With regard to this problem, the purpose of this paper is a rather modest one: first, arguing that it is a metaphilosophical problem and, second, suggesting that looking at it in this way may offer some useful insights about the metaphysical issues that are typically raised by the clash of the images. I will not venture any solution to the problem though, nor in fact will I address the questions that concerned Sellars himself the most about it.[1] The plan of the paper, instead, is the following. In Section 2, I will talk about what metaphilosophical problems are, how they do arise and why the clash of the images can be understood as rising one. Sections 3 and 4 are devoted to characterizing the different normative structures of substance concepts in the manifest and the scientific images. Section 5 rehearses the Sellarsian thesis of the fundamentality of the concept of a person in the categorial articulation of the manifest image, with the pur-

[1] In particular, I will not discuss the problem of the unity of the person and the accommodation of sense impressions in the scientific image.

philinq X, 1-2022, pp. 49-72
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.388
Submitted: March 2021
Accepted: October 2021

pose of distinguishing the ontological and the semantic issues involved in such a fundamentality. Eventually, Section 6 explores the problem of recategorizing normativity in the scientific image. Hopefully, this path will lead to a better awareness of what the actual means of philosophical investigation are with respect of the clash between the manifest image and the scientific image.

2. *Are there metaphilosophical problems?*

The status of metaphilosophy as a discipline is a controversial one. Intuitively speaking, metaphilosophy is supposed to be the philosophical investigation of philosophy itself, of its domain, methods, goals and general meaning as an intellectual endeavor. Those who may wish to reject the idea of a philosophical discipline whose task is to conduct such an investigation have a rather strong argument on their side, an argument purported to show that the very notion of metaphilosophy is somehow inconsistent. It goes more or less like this: The self-reflective stance required by metaphilosophical investigation can either be adopted outside philosophy or inside it. If the former, then metaphilosophy is not a *philosophical* discipline – it must rather pertain to some other field, like, perhaps, sociology or the history of thought. But if the latter, then metaphilosophy is just philosophy at large and so it is not a philosophical *discipline* – it just what philosophers do when they think "carefully", which is what they are always supposed to do.

The problem with this argument should be apparent. It assumes that metaphilosophy, as a discipline, is to be characterized in terms of a certain "sideways on" stance that those who practice it should adopt. Such a view is probably conveyed by the prefix "meta". In effect, it is the same idea epitomized by the way in which the notion of a *meta*-language is sometimes conceived, namely: Metalanguages are required because you can't talk about an object language by means of its own expressive resources, on pain of contradicting yourself. So, the idea is, just like you need to put yourself in a position external to a language in order to talk about it, similarly you need to put yourself in a position external to philosophy in order to talk about it. But that, of course, is just as wrong for metalanguages as for metaphilosophy.[2] And even more so, given that it is not so clear what the expressive resources of metaphilosophy are

[2] What Tarski's undecidability theorem shows is that we should pay attention to the way in which we use the expressive resources of our semantic vocabulary, especially the truth predicate (Tarski 1933; 1944). This result has some very important consequences on the model-theoretic metatheory of arithmetics, but it by no means implies that, in order to talk about an object language, we must pull ourselves out of it in a *different* linguistic practice. On the other hand, the lesson about the expressive power of our metavocabularies will come in handy below.

supposed to be and whether or not their use could engender contradictions.

So, the argument would show that metaphilosophy is inconsistent *if* metaphilosophy was a stance or an attitude, which, as we have just seen, is not necessarily the case. But what is the alternative? Timothy Williamson (2007) famously proposed to conceive metaphilosophy as "the philosophy of philosophy," where, of course, the genitive is to be construed as an objective one. In this sense, it does not specify as much an attitude of the philosopher as the object of study of the philosophical discipline, and "metaphilosophy" is just to be considered as a bad name for what it should refer to. Unfortunately, we are no more in the position of changing our terms, if we ever were. Williamson had as good chances as anyone to do it, but failed. Oddly enough, "metaphilosophy" must have had some "philosophical" ring in it and it sticked. Be that as it may, there is no reason now not to comply with the standard practice, once the meaning of the term is clear.

In Williamson's sense, it is an analytic truth that metaphilosophy is a philosophical discipline, just like it is for the philosophy of science or the philosophy of mind. This approach, however, may raise another familiar sort of puzzlement. There can be little doubt, if any, that these are all disciplines and that they can be distinguished from one another in terms of their object of study. But what is it that makes them *philosophical*? What is the difference, for instance, between the philosopher of mind and the neurobiologist who thinks carefully and "persistently reflects" about her discipline? These are precisely the sort of puzzlements that Sellars thought he could address and clarify by means of the famous definition of the aim of philosophy that he proposed in the first line of *Philosophy and the Scientific Image of Man*.

We will come back to the content and the implications of his proposal below. For the moment, it is more relevant to focus on the fact that the line of reasoning that we have just sketched exemplifies *one* of the ways in which metaphilosophical issues typically arise.[3] They do arise, in this first manner, when special sciences claim for themselves specific subject-matters that traditionally pertain to philosophy. When that happens, philosophers rarely retreat and abandon the disputed field to the newcomers, who, in general, tolerate them staying and simply go on with their own research and business. The problem araises for the philosophers instead, who need to justify their presence in the field that has been expropriated to them and is soon characterized and governed by different rules and practices. Thus, they may begin to ask in a

[3] Of course, philosophers have decided and may decide to engage in metaphilosophical reflections for the most various reasons. In the following, I will simply focus on two ways in which metaphilosophical issues may arise that are especially useful for the characterization of the problem raised by the clash of the images.

crescendo: What is the specific contribution of philosophy in this field? What is the specific contribution of philosophy to any field whatsoever? And, eventually: What is philosophy really about? In this first sense, then, metaphilosophical questions arise as a consequence of the demarcation of special domains of scientific investigation and their expropriation in favor of special disciplines. They arise out of the need for a reinterpretation of the philosophical contribution to the research in those domains, that was unreflectively taken for granted before.

This is probably the most common source of metaphilosophical issues, but it is neither the only or the most serious one. A second way in which metaphilosophical issues may arise occurs when the very categorial articulation that traditionally characterizes philosophical discourse[4] is challenged with regard to the representational adequacy of the theories that can be provided in terms of it.[5] Of course, this is the way in which the clash of the images is relevant to metaphilosophy. It is essential not to be mistaken about what "categorial articulation" is supposed to mean. For the moment, let us be content with Sellars's suggestion that the best way to start thinking of categories is to see them as "*summa genera* of conceptual items" (Sellars 1970: §24), or, better, "classifications of conceptual roles" (Sellars 1981: II, §81). Categories will be further explored in the next sections, but a few remarks are still in order here about what the issues generated by challenging their representational adequacy are not.

From a formal point of view, representational adequacy is best conceived as the possibility to define homomorphisms between a representing structure and the represented one. For our purposes, an homomorphism can be thought just as a mapping between the structures that preserves the properties of their elements and the relations between them. So, for instance, if I want to represent married couples as sets of points, my representation can be said to be adequate in this sense if, whenever it is true that *A* and *B* are married, then it is also true that the points into which *A* and *B* are mapped belong to the same set. More interestingly and generally, my conceptual representations can be said to be adequate in this sense if states of affairs can be homomorphically

[4] Here "discourse" is intended in a rather noncommittal sense as referring comprehensively to the expressive resources by means of which philosophy is practiced. The most significant parts of this discourse will be more thoroughly discussed below.

[5] At this point, already, some of those who even broadly sympathize with the Sellarsian enterprise, but harshly oppose representationalism (as e.g. the Rortian pragmatists) might be willing to jump off the train. My recommendation is to resist anti-representationalism at this early stage of the argument, at least for the sake of it. Even if one is not really keen to embrace Sellars's own "picturing" idea in order to make sense of the representational adequacy of different conceptual structures (Sellars 1968: ch. V), there are alternatives. Brandom and McDowell, pre-eminently, have shown how to make use of representational semantics while keeping its epistemic implications at bay (cf. Brandom 2013).

mapped into them.

Now, one reason why representations may fail to be so adequate is because the properties and relations of the representing structure are not sufficient, in number or type, to preserve those of the represented one. Thus, one could think that challenging the representational adequacy of the categorial articulation of philosophical discourse may consist in providing a better, or just different conceptual repertoire.[6] That, however, is a mistake. Categories are not conceptual repertoires: categories classify concepts of the same sort. So, different sets of concepts can be accommodated within the same set of categories, or, more precisely, different conceptual repertoires can share the same categorial articulation. By way of example, consider the difference between the concept of *miasma* and the concept of *virus*. They clearly belong to different conceptual repertoires: the former to the ancient Hippocratic medicine, the latter to the modern one. Both repertoires, however, share (part of) the same categorial articulation. In particular, both the concept of miasma and the concept of virus are *substances* that *cause* certain diseases. Of course, the problem could be raised of which one of them better "carves at the joints".[7]

But that is precisely a question of the sort that one could take to be properly addressed *in* philosophy.[8] Even if a question like that could foster some debates about the constitution of specific disciplines, like ontology or semantics, it has no direct impact on the way in which philosophy itself is construed, nor does it imply reorienting what philosophers do.[9]

Challenging categories is something else entirely. When the representational adequacy of a categorial articulation is put into question, the adequacy is

[6] A conceptual repertoire can be thought as an inferentially articulated collection of concepts in terms of which representational structures can be constructed. As far as I can see, in Chapter V of *Science and Metaphysics* Sellars adopts the term "conceptual structure" for the *cognate* notion of a (cross-)linguistic repertoire of resources to picture things that is available at a certain stage of the development of linguistic practices.

[7] In the sense of Lewis 1983 or Sider 2011.

[8] Needless to say, one could insist instead that the task of providing the right conceptual repertoire for representing how things are pertains to natural sciences. Or, one could even argue that the question itself is a bad one, because the choice between different conceptual repertoires is ultimately indeterminate with respect to the states of affairs that they are used to represent. And by no means these are the only views that could be endorsed in this debate. But the point here is just that they are all philosophical views and the debate is a philosophical one.

[9] Admittedly it *could* have a metaphilosophical impact, yet an indirect one. Thus, one could think that questions of that sort cannot have any real answer and, therefore, that trying answering them is not what philosophy should do. Notice, however, that this metaphilosophical view is not motivated as much by the question about what conceptual repertoire is better as by the realization that no conceptual repertoire is representationally adequate. In this sense, this is an instance of the second way of generating metaphilosophical issues that we are just about to characterize.

contested of any representation developed by means of *any* conceptual repertoire that shares such an articulation. And that is far more radical a challenge. As far as the topic of this section is concerned, it implies that the very questions about what concepts better carve at the joints should not be addressed in the same way as before and, therefore, that the philosophical discourse providing the context for the debate about such questions should change. This is precisely the second way in which metaphilosophical issues may arise. In this sense, they arise when something wrong is identified in the philosophical discourse, something that requires a revision of the way in which philosophy is thought and done.

In this section, I have provided a concise characterization of metaphilosophical problems and their origins. With no claim of being exhaustive, I argued that there are at least two ways in which they can arise. In both cases, they do as a consequence of a challenge that puts the usual philosophical practices into question. Such a challenge can be of at least two kinds. It can be the more narrow one of contending for specific subject matters with other disciplines. Or, it can be the broader one of revising the categorial structure of philosophical discourse altogether. As anticipated, the challenge brought by the clash of the images is of the latter sort. So, that is what the next sections are going to deal with.

3. *Substances and powers*

When Willem deVries presents the notions of the manifest image and the scientific image for the first time in his introduction to Sellars's philosophy, he feels like warning the reader that the distinction has outgrown the domain that it was originally intended to pertain to and "has taken a life on its own" (deVries 2005: 9). What he has in mind is the fact that many of those who get acquainted with the Sellarsian distinction in the context of the debate on scientific realism are likely to understand it as depicting the categorial incompatibility between the way in which one and the same world is described by common sense and by science.[10] Thus conceived, the distinction would be nothing more than a philosophical gloss on Eddington's "two tables" example (Eddington 1928: ix). The core of the problem would be all already there. On the one hand, common sense says that a table is a white and wooden object. On the other hand, science says that the very same table is a collection of rapidly moving particles, deflecting photons and hitting with each other according to certain probabilities. Thus, the manifest image would be the image of manifest objects, like shoes

[10] For a classical use of the Sellarsian distinction in this diverted sense, see e.g. van Fraassen 1980.

and ships, while the scientific image would be the image of scientific objects, like particles and fields. If we think of the images this way, we will also be led to construe their clash as the conflict between the commonsensical reluctance to endorse the ontology of scientific theories as real and the scientist claim that the objects of common sense are merely apparent.

It is important to see why this interpretation of Sellars's original distinction is mistaken, even though the ontologies of the manifest image and the scientific image *are* indeed incompatible. As a matter of fact, the mistake is twofold. First, the real *locus* of the conflict is misplaced: the incompatibility between the images primarily lies in the logical articulation of their categories rather than the ontological commitments that the application of the conceptual repertoires articulated by those categories imply. Second, there's a misunderstanding about the theoretical purpose with which the distinction of the images was originally devised. Sellars aimed at problematizing the way in which philosophy is called by natural sciences to a recategorization of the discourse in which it has traditionally pursued its ethical and theoretical investigations. As we will see, this double misinterpretation prevents understanding the implications of the fact that the category of a person is fundamental to the metaphysics of the manifest image.

Admittedly, Sellars's own presentation in his 1962 essay does not really provide a clear understanding of the clash of the images beyond the ways in which it shows itself on the ontological level.[11] He stipulates that the scientific image is to be distinguished from the manifest one by the fact that only the former involves the postulation of imperceptible entities. That can't be right though, not until postulational methods are better qualified. Indeed, why can't manifest objects be postulated as well? Consider the following case: "I hear scratching in the wall, the patter of little feet at midnight, my cheese disappears – and I infer that a mouse has come to live with me" (van Fraassen 1980: 19-20). This seems just a sound instance of an explanation by postulational means, but it definitely comes out in favor of the continuity between science and common sense. Surely, one would rightly be reluctant to treat mice as postulated objects.

[11] Chapter V of *Science and Metaphysics* is tuned in the same ontological key. There, Sellars's focus is on the epistemic import of matter-of-factual statements and his aim is to explain how the objects pictured by a conceptual structure in the manifest image are "phenomenal" by showing that they can only be said to *really* exist if the concepts in terms of which they are identified have counterparts in the ultimate, "Peircean" conceptual structure in the scientific image. This way of illuminating the clash of the images, however, obscures the role of the categorial articulation of the different conceptual structures. In fact, the notion of picturing does not stand in place of, but presupposes an analysis of the semantic uniformities establishing the correlation in which a method of projection consists in (Sellars 1968: §58). And the way in which such semantic uniformities are determined is precisely what is made explicit in terms of a categorial articulation.

However, one should also be wary of reacting to this by insisting on the imperceptibility of postulated entities. Indeed, are the atoms postulated to explain why matter behaves the ways it does more "scientific" than the God of Thunder postulated to explain why it rains? Of course they are, but not because they are less perceptible. A more serious proposal along these lines would be to try and draw the distinction in terms of observational and theoretical objects. Unfortunately, for the present purposes, such a strategy would work only on condition of endorsing a crude empiricist notion of observation, according to which the content of experience is given independently of any conceptual mediation. For, as soon as the role of conceptual frameworks is acknowledged in the determination of the content of experience, the need would rise in turn for a distinction of the manifest from the scientific framework in order to separate observational reports pertaining to common sense and those pertaining to science. It is evident, therefore, that the strategy cannot be applied in the interpretation of Sellars, who is most famous for having attacked precisely this notion of observation as "The Myth of the Given".[12]

These preliminary remarks may well be trivial, but they are enough to show that the source of the clash of the images can not be traced back primarily, but only derivatively to the ontological level.[13] A far more insightful characterization of the differences between the scientific image and the manifest image is provided by comparing the postulational methods of the former with the correlational techniques of the latter. Sellars's favorite example to illustrate this contrast is the way in which thermodynamics explains the correlations between pressure, volume and temperature expressed by the law of ideal gases. For instance, one of such correlations (with a bit of simplification) is the following:

[12] This is not to say that the distinction between observational and theoretical objects does not make sense in a Sellarsian context. For the purpose of clarifying this point, I suggest considering an object theoretical if it is just not an observational one. The latter, in turn, can be conceived as an object to which reference can be made in observational reports. Thus, common sense observation can be intended as the perceptual experience of the layman, rather than the specialized and heavily theory-laden practice of the scientist in the laboratory. The latter is the sense in which, for instance, it is correct to say that Higgs bosons have been observed by the ATLAS and CMD detectors at the Large Hadron Collider at CERN in Geneva. This definition of theoretical objects might sound circular, but it doesn't have to. It is, however, a practical difference, in the sense that whether certain objects are treated as theoretical or observational in a certain discursive practice depends on the characteristics of the practice itself – and practices are not necessarily defined in terms of theoreticality. This approach is Sellarsian at least *in spiritu*: as deVries noticed, "[f]or Sellars, the observation-theory distinction is fundamentally a methodological distinction with no direct ontological import" (deVries 2005: 155), since nothing really prevents us from learning to make observational reports of theoretical objects.

[13] Notice that if it could, then the clash of the images would be a meta-ontological, rather than a meta-philosophical problem.

1. If the temperature of a gas increases, then its pressure increases as well.

Now, the kinetic theory of gases accounts for (1) by postulating imperceptible particles, whose existence corresponds to the existence of the gas, whose average kinetic energy corresponds to the temperature of the gas and whose hits on the walls of the container in which they are placed correspond to the pressure of the gas. Given correspondence rules roughly like these, the empirical generalization formulated in (1) can be explained in terms of a theoretical generalization such as:

2. Faster particles likely hit the walls of their container more frequently.

Sellars argues that theoretical generalizations like (2) explain why manifest objects obey empirical generalizations by means of the correspondence rules that identify them with (collections of) theoretical entities (Sellars 1961). Both (1) and (2) are lawlike generalizations that could be formulated as universally quantified conditionals, but – and this is the crucial point – they do not express laws of the same sort. Thinking otherwise would be failing to appreciate the categorial difference between the manifest and the scientific image.

Let us focus on (1) first. So, (1) expresses a correlation between two properties of the gas: being heated to a certain temperature and being under a certain pressure. Sure, temperature and pressure are to be considered quantitative concepts in their own right (cf. Carnap 1950). But the fact that the application of the concepts "being hot" and "being compressed" have been subjected to measurement in certain practices should not mislead us – no more than the fact that the truth of "Socrates is pale" could become a matter of degree on a scale devised by tan fanatics. Manifest objects have properties that regularly change in correlation with other properties. Just like Socrates becomes more tanned if exposed to the sun in standard conditions, so gasses become more compressed if heated in standard conditions. This much is what the empirical generalization expressed by (1) is about. But why is it lawlike?

Sellars's most comprehensive discussion of laws of nature in the manifest image can be found in *Counterfactuals, Dipositions and the Causal Modalities* (1957). There are two results of his analysis that are the most relevant for the present purposes. First, he argues that causal relations are expressed by counterfactually robust conditionals, rather than mere accidental generalizations, because modal vocabulary makes explicit the endorsement of defeasible rules for inferring from the assertion of the causes to the assertion of the effects. Second, our concepts of manifest thing-kinds are constituted by the rules of inference that are made explicit in terms of the subjunctive conditionals expressing causal relations.

The first thesis, of course, is testimony to Sellars's inferentialism, according to which conceptual contents are determined by their inferential articulation. Surely, there is a problem with the idea that laws of nature are expressed by counterfactual conditionals, if it is construed as condition for identifying causal relations, because not all the generalizations underwritten by counterfactually robust inferences are laws of nature – e.g. "If this apple were red, it would be colored" underwrites the semantic generalization that red things are colored – or laws at all – e.g. "If I were to choose a coin at random from my pocket, it would be copper" underwrites the accidental generalization that all coins in my pocket are copper (Brandom 2008: 105). Fortunately, we don't need to adjudicate here whether or not Sellars really endorses this idea. We will be content with a weaker reading of the thesis, according to which laws of nature, among other generalizations, are expressed by counterfactually robust inferences.

The second thesis grounds our understanding of substance concepts on the empirical generalizations that are underwritten by laws of nature. In the manifest image, the latter explain what things do when acted upon in certain circumstances. That is why the "logical form", as Sellars put it, of an empirical generalization in the manifest image is "*K*s φ when ψed in *C*". According to Sellars, this specific kind of empirical knowledge is embodied in substance concepts and sets the criteria for their application apart from those of the other concepts – in particular, it is what allows us to use them to re-identify things and stuffs in the world. In this sense, manifest substances are things endowed with dispositional ("iffy") properties, or "powers". The notion of causality in the manifest image is essentially intertwined with the conception of such powers.

These two theses together contribute to characterizing substances in the manifest image as having powers that determine their identity conditions and articulate the causal relations between them.

4. *Changing categories*

Let us now turn to (2) and try to see how it is supported by a different categorial structure, a different notion of substance and a different notion of cause. Surely, (2) is in striking contrast to (1) from an ontological point of view, given that it makes reference to a different sort of objects: the micro-particles postulated by the kinetic molecular theory. But how are these micro-particles really different from gases, i.e. the manifest objects that they are supposed to correspond to? We have already ruled out that the question could be answered by focusing on epistemic distinctions. A more promising suggestion would be then to draw the line directly by ontological means. So, for instance, one could

try and notice that while molecules are parts of gasses, they have no parts in turn. And even if they do, still there must be some ontological "atoms" that will explain the dispositional properties of all the less fundamental entities up to manifest objects like gasses.[14] Although there is something to this suggestion, it is ultimately wrong-headed in this case. The reason why these "atoms" would be suitable to play the explanatory role that Sellars ascribes to the postulated entities of micro-theories does not depend as much on the fact that they are ontologically fundamental as on the fact that they do not have powers. If they did, a more explanatorily fundamental level of scientific entities would be required in order to account for the manifest correlations determined by such dispositional properties. Once this is clear, it is easy to see that ontological fundamentality is neither sufficient nor, in fact, necessary for the sort of explanation that Sellars has in mind for the scientific image.[15]

Conceiving the particles of the kinetic theory of gases as scientific objects then is, primarily, not categorizing them in terms of concepts for manifest thing-kinds. This means that it is a mistake to read (2) as "Particles hit the walls of the container more frequently when accelerated" and as expressing a causal property of the particles. The mistake consists in overimposing the categories of the manifest image on the concepts of the scientific one. A mistake that can be amended only by understanding the criteria with which the latter are applied and articulate empirical knowledge. Just like we did for (1), then, it is worth starting by asking why (2) is a lawlike generalization.

It is tempting at this point to be led to emphasize the fact that at the micro level "in the fundamental laws of physics there are odds" (Feynman 1967: 145). Indeed, it could be argued that the probabilistic character of the laws of quantum mechanics is what makes them most puzzling for common sense. And yet, indeterminism is a red herring as far as the problem of understanding the categorial distinction between the manifest and the scientific image is

[14] Granted that wholes might have dispositional properties that none of their parts have, still an explanation must be provided of how such powers emerge given their ontological constitution. The burden of the proof is on those who deny that some sort of such an explanation is required.

[15] This is how I understand Sellars's observation that micro-theories postulating micro-thing-kinds with the same logic of manifest thing-kinds do not take us all the way to the scientific image (Sellars 1957: §51). This point must be carefully distinguished from the gist of Sellars's reiterated remark that "our work-a-day descriptions of the [conceptual] episode are [...] of a mongrel hypothetical-categorial character" (Sellars 1968: VI, §4). In particular, the idea that in the manifest image we can provide only a functional classification of our inner episodes in analogy with the *ought-to-be*s (semantic uniformities) governing our overt linguistic behavior, and that in the scientific image we will have at our disposal determinate conceptions of the categorical character of such episodes does not imply *per se* that our descriptions of such episodes in the scientific image will do without mongrel hypothetical-categorical statements.

concerned. In order to see why, we need to distance ourselves a bit from the analysis that Sellars gives of probability in the context of establishing empirical generalizations.

To begin with, Sellars notices that stochastic methods are part of the tools that scientists use in the manifest image as well. This use of probability is not really problematic for our purposes. It applies to cases in which a system is so complex that we just cannot know everything we would need in order to calculate the outcomes precisely and so we attribute a certain probability to each of them. For instance, if I were to roll a 6-sided unloaded dice, I could, in principle, predict how it will land by means of the purely deterministic laws of classical mechanics. But knowing all the variables required to make such a prediction is practically impossible, so we prefer to describe the system by saying that the probability of rolling each number is 1/6 (*ibid.*). Another example is again the way in which statistical mechanics is applied in the kinetic theory of gases. In this sense, stochastic methods are an expressive resource that can be accommodated within the manifest image because it does not imply any change in the logic of empirical concepts. Indeed, according to Sellars, the purpose of probability in inductive inference is to make explicit that there are empirical reason to make certain assertions (Sellars 1954: §§57-71; 1957: §60; 1964).

The sense in which probability is used instead within the very laws of physics to express the fact that nature is intrinsically nondeterministic seems to go deeper into the structure of our empirical knowledge. And in fact it does, but it is important not to run together the implications of the nondeterministic conceptual repertoire provided by quantum mechanics with the implications of the scientific categorial structure to which it belongs. While the latter are what we are trying to investigate here, the former could in principle be the same also for a conceptual repertoire in the manifest image. Suppose, for instance, that dices were intrinsically nondeterministic thing-kinds, to be understood (in Sellars's account) in terms of dispositional properties like "Dices, when rolled, land on 5 with a probability of 1/6".[16] Of course, the laws so expressed would be probabilistic and that would have implications for our concept of causality, but the reason why such generalizations would be lawlike is just the same as we have seen before. In other words, stochastic powers would still be powers.

Particles, *qua* scientific objects, have no powers. But what does that mean? Actually, the sense in which entities in the scientific image have no causal properties is not so hard to understand. And in fact, it can be grasped even without mobilizing the counterintuitive framework of quantum mechanics. For the sake of the argument, however, let us assume that physics provides in effect the

[16] Would that really sound so bizarre to the layman?

paradigm of a scientific representation of the world of nature. Physicists are trained to study physical systems and to apply mathematical tools to represent them. A complete description of the state of a physical system can be provided by specifying the values of the *n* variables that represent the properties of the system that change with time.[17] All the possible states of the system can be represented as points in a *n*-dimensional space, the so called "phase space" of the system.[18] A series of equations are then defined to represent the evolution of the system in time. By filling in all the information required to specify a present state of the system, a physicist can determine the behavior of the system in the future *and* in the past.

The particles composing a gas in the kinetic theory are a physical system, whose behavior is understood along these lines. When we say that in the scientific image "a gas is [...] a cloud of molecules" (Sellars 1961: §41), what we mean is that a gas is a physical system. Notice, however, that there do not really seem to be "objects" in a physical system. Of course, as long as physical systems are studied of which we also have more intuitive representations in the manifest image, it is easy to treat them as merely useful technical tools to make calculations and predictions about *things* that do not really have such a categorial structure. But the clash of the images is already there. In the case of the systems of quantum mechanics it is just more painful, because no manifest intuition is available to challenge the reality of the mathematical representations.

There do not really seem to be "causes" either in the evolution a physical system. Surely, there are no causal relations between *doings* and *results* of the sort underwritten by the powers of manifest thing-kinds, for the simple reason that there are no such kinds in physical systems. But there are also other aspects of the intuitive notion of causality that are lost in the mathematical representation of a physical system. The most striking one is the absence of any definite sense in which in a physical system the effects should be construed as following and be determined by the causes. In fact, given a certain state of the system the very same equations can be used to calculate a state in the future as well as a state in the past. More generally, if the scientific image is acknowledged as real, it is not clear how the calculation of the evolution of a system could be construed as the representation of a given state causally determining another one. It is precisely along these lines that Russell (1913) went on arguing

[17] So, for instance, in order to describe the state of the system of a particle moving on a line, the values of just two variables must be specified: the position of the particle in one dimension and its momentum.

[18] More precisely, the phase space is a vector space in which a state of the system is represented as the coordinates of a vector specifying the properties that change with time. Since we will not do any calculation, however, we don't really need the mathematics of vector spaces either.

for his famous dismissal of the law of causality.[19]

Are there "laws" that govern the behavior of a physical system? This is of course how the question with which we began the analysis in this section has to be rephrased now. One might be tempted to answer straightforwardly in the affirmative: what else the equations with which the evolution of the system is calculated would express otherwise? As a matter of fact, however, the point is a little more delicate than that. On the one hand, it is correct to say that the equations are not underwritten by contingent regularities.[20] When a physicist applies the equations to calculate the evolution of a physical system, she does not merely mean that the system could have been in a given state in the past and could be in a given state in the future, but that it *was* (with a certain probability) in a given state in the past and *will be* (with a certain probability) in a given state in the future. On the other hand, however, it is not equally clear that she means that the system is *governed* by the laws that are expressed by the equations or that the evolution of the system *abides by* them.

Let us try and sharpen our intuitions about this distinction by thinking about one more question (in the manifest image).[21] Suppose that *all* (all!) *possible A*s are also *B*s: Is that enough to say that it is a *law* that all *A*s are *B*s? On the one hand, one might be willing to answer that the absence of possible (not merely actual) counterexamples is indeed enough to make sense of a universal generalization as a law. In this sense, the fact that all possible *A*s are *B*s is the reason why it is correct to infer that something is *B* under the assumption that it is *A*. This first intuition amounts to the idea that the laws of nature are not really normative: they are just general facts. Norms govern only our inferential practices in view of the facts that we aim to represent by means of them. On the other hand, one might be willing to answer in the negative instead and rebut that the notion of the absence of counterexamples is just different from the notion of a normative connection in which a law consists. This second intuition has it that the modal vocabulary in which the former notion is formulated is but an expressive resource to make explicit the latter in the "material mode".[22] It is not hard to see that the distinction between these two intuitions tends to

[19] He was, of course, followed by Wittgenstein, who, in the *Tractatus*, denied the existence of a "causal nexus" (5.136) and at the same time – more insightfully – acknowledged that causality is not a law but "the form of a law" (6.31).

[20] In view of the general Sellarsian framework in which the present discussion is conducted, it is legitimate enough, I reckon, to assume this without argument.

[21] This is the very same distinction that Sellars discusses by means of his counterfactual example about planets revolving around a central sun (Sellars 1957: §65).

[22] Sellars famously noticed in this regard that "the language of modality is [...] a 'transposed' language of norms" (Sellars 1953: 332).

be congruent with the distinction between the notion of a law in the scientific image and the manifest image respectively.

This last observation allows us to introduce one last point that is crucial to make before moving on to the next section. It has to do with the fact that the terms "objects", "causes" and "laws" were put in scare quotes above when we questioned their *existence*. The reason for adopting this typographic practice is that those terms do not really refer to *things* that could or could not exist in the world.[23] They are examples instead of the categories that articulate the structure of a conceptual framework. As was pointed out, according to Sellars they provide classifications of conceptual items, i.e. they are terms that belong to the syntactic metalanguage of rules and are applied to make explicit how the expressions of the object language are used. In this sense, taking the questions about their existence as they are formulated *in the material mode* at face value is a mistake. When they are transposed *in the formal mode*, such questions ask about the empirical reasons for the use of referring expressions, the application of substance concepts and the assertion of the conclusion of inferences. Saying that the conceptual frameworks of the scientific image and the manifest image have a different categorial articulation is saying that such questions receive different answers with respect to the two frameworks. Therefore, it is a mistake to say that there are no objects, causes or laws in the scientific image. The truth is that those are categories that change in the scientific image.[24]

5. *The fundamentality of persons in the manifest image*

On numerous occasions Sellars insists that the concept of a person is fundamentally intertwined with the categories of the manifest image. At first glance, the ties seem to go both ways. On the one hand, people are used to conceive of themselves in terms of the manifest image as intentional subjects and rational agents. On the other hand, the manifest image can be said to be the conceptual framework of persons, because its categorial articulation is modelled right after the concept of a person. According to Sellars, however, the first direction of dependence is merely contingent, in the sense that people can also conceive of themselves in terms of the scientific image. As is well known, he argues that the concept of a person is not essentially grounded in the manifest image and can be eventually recategorized as a system of scientific objects (Sellars 1969:

[23] This idea is just the core of the "rationalistic metaphysics" that is attacked in this regard by Sellars (1957: §106; 1970: §§17-31) – and by Wittgenstein.

[24] Of course, *in the material mode one* could still say that in the scientific image there are no *manifest* objects, no *manifest* causes and no *manifest* laws, but there are *scientific* objects, *scientific* causes and *scientific* laws.

§59) – possibly, a bundle of absolute processes (Sellars 1981: III, §125). This conclusion, of course, is far from obvious and securing it is what Sellars is most interested in. As far as our present purposes are concerned, however, the dependence of the manifest categories on the concept of a person is far more relevant. By way of clarifying such a foundation, Sellars puts forward two lines of reasoning that we will briefly rehearse here.

The first one is framed as a genealogy. According to it, the manifest image is a refinement of an "original" image. Just like the manifest and the scientific, so the original image too was an all-encompassing conceptual framework, but its most significant characteristic was the fact that *all* the objects in it were persons: not only men and women, but every *thing* was conceived in terms of the concept of a person as an intentional subject and a rational agent. This means that, in the original image, rivers, for instance, flow towards the sea because they intend or are used to do so, and if someone were to divert their course they could get angry, so that they should be prayed to convince them not to overflow. In other words and more straight to the point, in the original image every event is explained in terms of the paradigm of intentional actions, either as the result of an object's intentions or habits. Therefore, the only sense for an episode to be caused in the original image is the sense in which a person causes another to do something that she wouldn't otherwise do (Sellars 1962: 13).

The manifest image would have emerged from this original conceptual framework by means of "a gradual pruning of the implications of saying with respect to what *we* would call an inanimate object, that it *did* something" (Sellars 1962: 12), so that the only sense in which such objects are expected to do something is by habit, disposition or power. When Sellars says that the manifest image results as a *refinement* of original one, he means that this "pruning" does not amount to a modification of the categorial structure of the image. This is important and should be acknowledged at face value: *all* the objects of the manifest image are still persons – although some of them are "truncated" ones.

Now, of course, the historical plausibility of a panpsychist conceptual framework that could ground the reality of the original image is not to the point here, because the genealogy proposed by Sellars is not to be intended as an argument to prove the dependence of the manifest image on the concept of a person. Sellars's purpose is rather to articulate the sense in which the manifest image is the conceptual framework of persons. Thus, the genealogy makes clear that in the manifest image substances are shaped upon the concept of person and why causality is conceived in terms of their dispositional properties.

The second and more important line of reasoning that helps clarifying the dependence of the manifest image on the concept of a person is developed by Sellars with relation to the normative account of the determination

of conceptual contents. Indeed, according to Sellars, conceptual contents are determined by the rules that govern their application in the space of reasons. In the absence of such rules no conceptual framework could be conceived and the question of its representational adequacy would not even make sense. Now, Sellars argues that *in the manifest image* the normativity that is constitutive of conceptual contents is grounded on the concept of a person. In brief, the reason is that only persons are suitable to play the game of giving and asking for reasons. But the details of Sellars's argument are so much articulated and delicate that it is often tempting to take a shortcut through some of them, with the risk of concluding something more or something different than what Sellars originally intended to. Since there is no space here to follow the argument through all its steps across the different texts in which Sellars unfolds it, I have to take that risk myself and provide but an outline of its main parts.

The first part is the idea that we can primarily understand the semantic content of our episodes of conceptual thinking only on the model of the pragmatic significance of our episodes of outer linguistic behavior. Such a pragmatic significance is articulated in terms of *ought-to-be*s to which our spontaneous dispositions to think-out-loud must conform. The second part is the idea that these *ought-to-be*s imply *ought-to-do*s, in the sense that the reason why our linguistic behavior is not merely regular, but rule-governed is because we have been taught to exhibit it by adults who intentionally acted so that we conformed to the norms that determine the pragmatic significance of our languagings. The third part is the idea that normativity has a transcendental condition: we are properly bound by rules to the extent that we share intersubjective *we*-intentions that rationally motivate our actions as members of the same community. Since persons are what can share *we*-intentions, they are fundamental to the possibility of determining conceptual contents in the manifest image.

Two lines of reasoning have been rehearsed here to explain why persons are fundamental in the manifest image. The first one sheds light on the fact that the manifest category of substance is shaped upon the concept of a person, in the sense that all manifest objects are either persons or "truncated" ones. The second one reconstructs the role of the concept of a person in grounding the normativity of conceptual contents. It is essential not to run the two lines of reasoning together. The first line of reasoning deals with an ontological issue, the second line of reasoning addresses a semantic issue.

6. *Norms in the scientific image*

In the opening of this paper the clash of the images was described as the problem of whether or not philosophy is entitled to deal with the human be-

ing by its own means. Apparently, the idea that special sciences have successfully competed with philosophy for more and more specific subject-matters and have imposed the categories for their investigation implies that philosophy is ultimately an alien in the scientific image and that it should be at home only in the manifest image. It is now time to zero in on such a presupposition and consider what the proper means of philosophy really are.

Sellars' own solution to the clash of the images is notoriously problematic. On the one hand, one of the most exploited "ideal types" in his rhetoric is the notion of a "perennial philosophy", which endorses the manifest image as real. Indeed, the idea of a perennial philosophy implies that a significant part of at least the western philosophical tradition is in no position to work with the categories of the scientific image. On the other hand, of course, Sellars thinks that philosophers should appreciate the picturing dimension of conceptual structures and learn to navigate the categorial changes leading through the scientific image to the Peirceian conceptual structure, in terms of which ideally adequate representations can be provided (Sellars 1968: V, §§72-74). Yet, when it comes to envisaging a strategy for resolving the incompatibilities generated by the clash of the images, it seems that the best he can offer is a metaphor and an advise. The metaphor of a stereoscopic vision in which the images can be brought together and the advise not to rush and endorse the scientific image as real in its current categorial articulation, which is presumably not yet complete. This is hardly enough to cash out as a metaphilosophical thesis the claim that "the aim of philosophy [...] is to understand how things in the broadest possible sense of the term hang together in the broadest possible sense of the term" (Sellars 1962). This, however, is also a quite ungenerous way of framing Sellars's contribution.

I will not attempt to provide here a full blown interpretation of his metaphor and the solution to the clash of the images that he sought to distill thereby. My purpose is the different one of bringing to light the metaphilosophical implications of the clash itself. In order to do that, however, I need to address at least a possible misunderstanding lurking in the characterization of the scientific image as a personless conceptual structure. The misunderstanding is generated precisely by confusing the ontological sense and the semantic sense in which the concept of a person is fundamental to the manifest image that were described in the previous section. Such a confusion could bring forth an argument like the following one: The ontological fundamentality of persons is required to ground normatively determined conceptual contents, *therefore* conceptual contents cannot be accommodated in the scientific image.

If such an argument were sound, one could actually use it to attack the very notion of the scientific image as a conceptual framework as an incoherent

one.[25] As it is easy to see, however, the argument has a few things wrong with it. In order to highlight the problematic points that are most relevant to us, it is worth unpacking it a bit and make explicit three main ideas that support it:

A. Conceptual contents are constituted by norms.
B. Norms are grounded on the shared intentions of discursive practitioners.
C. Persons do not belong to the scientific image.

Of course, generally speaking, all these ideas are highly controversial. For the sake of the present purposes, however, let us look at them mainly from a Sellarsian point of view. If we do that, we can't but accept (A), because it is the cornerstone of Sellars's semantic theorizing.[26] The other two ideas are more problematic.

Let us consider (C) first. According to Sellars the claim that the ontology of the scientific image does not include persons is simply mistaken. The correct claim would be that while scientific substances are not persons, it is not obvious that manifest persons could not be recategorized in the scientific image as pluralities of scientific objects. For such a recategorization to be adequate, it must also take into account those social features that allow normativity to be grounded (Sellars 1962: sec. VII). Indeed, recategorizing the concept of a person in the scientific image is arguably one the core tasks of Sellars's whole philosophical enterprise.[27] Here, we are not interested as much in his results as in his methods. Sellars did not conceive recategorization as the task of providing the genera of empirical concepts whose species will compose conceptual repertoires in the scientific image. That is not the task as much of philosophy as of empirical sciences. What he did was making explicit the articulation of the manifest concept of a person in terms of the normative relations in which the concept is embedded in the manifest image. Such an articulation provides the criteria for recategorizing the manifest concept of a person, in the sense that the structure of those normative relations has to be

[25] In effect, Brandom (2015) argues that Sellars's whole distinction between a manifest and a scientific image of man in the world is a "misplaced [...] attempt to naturalize Kant's transcendental distinction between phenomena and noumena" in view of his scientific realism, that threatens the possibility to treasure Kant's fundamental semantic insight that conceptual contents are to be understood in normative terms.

[26] I will not argue for this thesis here. I just want to point to the fact that rejecting the normativity of conceptual contents simply cuts the Gordian knot of the clash of the images. Such an approach could well quieten the puzzlement generated by the problem, but at the price of ignoring its intricacies. That would be unfortunate, because there might be some lesson to be learned there.

[27] He mostly pursued it with respect to intentional episodes – in particular, of course, immediate sense experiences. Unfortunately he did not elaborate as much on the scientific successor-notion of collective intentions.

respected by the conceptual repertoires that will succeed in the categories of the scientific image.

Now, if (C) is dropped, as Sellars explicitly recommends, clearly the argument does not go through anymore. That is to say, the conclusion does not follow that the notion of the scientific image is incoherent. However, while (A) and (B) are in place, the sense in which conceptual contents would result to be accommodated in the scientific image is still problematic. Indeed, if persons are accepted in the ontology of the scientific image only as derivative from ultimate entities, the normativity that their discursive practices are supposed to ground would make sense only as *ideal* (in the Kantian sense). This disappointing result that seems to force upon us the mere ideality of our "second nature" has been variously resisted in the Sellarsian scholarship by proposing different interpretations of the synoptic view promised by the metaphor of the stereoscopic vision. A discussion of such proposals, however, is beyond the purposes of this paper. Especially since the third of the main ideas that support our argument against the scientific image is still to be addressed.

So, are there good reason to abandon (B)? Many have thought so. According to Millikan (2016), for instance, there are indeed good reasons, but they lead away from Sellars's original path, so much that she thought that having followed them she deserved the epithet of a "renegade daughter" of his. As is well known, she developed an account of (some of) substance concepts as naturally and culturally selected cognitive abilities to keep track of things in the world. Her approach to the naturalization of the space of reason, however, is not the only one. She pursues the task of explaining how conceptual contents are determined in terms of the selective advantage of the organisms possessing the corresponding cognitive abilities. An alternative to this piecemeal strategy, so to speak, would be to conceive the effects of the evolutive pressure on our representational abilities not primarily at the level of the processes that put our cognitive systems in factual relations with things in the world, but at the level of the cognitive abilities that enable us to engage in social practices (cf. Rouse 2015).

Of course, the mere existence of alternatives is not good enough a reason to reject (B). However, since our purpose is not as much defending the scientific image as learning something about it, considering possible paths for the naturalization of normativity is extremely useful, not only because they offer another opportunity to reason about the categorial differences between the conceptual structures of the manifest image and the scientific image, but also because in this case the clash does not reveal itself in the ontology and so one does not run the risk of being misled by ontological issues. In this case the question is not about the scientific successor of a given manifest concept, but about the scientific successor of the normative structure that allows making

sense of the categorial articulation in the manifest image. However, just like the question whether or not manifest persons exist in the scientific image does not do justice to the task of recategorizing the concept of a person, so lamenting that conceptual contents in the scientific image are not determined in terms of manifest normativity is missing the point. That might well imply that one is better off by saying that one does not really recognize conceptual contents in the scientific image anymore, just like one does not recognize persons in it. And it might also imply that telling what normativity is in the scientific image turns out to be, ultimately, an empirical task. Sellars's advise not to embrace the scientific image in its current form, at this point, seems to be as sensible as ever. Even at this level, however, I suggest that the task of the philosopher does not really change. In the case of the recategorization of a concept, she has to develop the expressive resources to make explicit the rules that constitute its content and to keep track of it across different conceptual structures.[28] In the case of normativity, she must develop the expressive resources to make explicit not only the rules by which we play the game of giving and asking for reasons, but also keep track of them across different conceptual structures.

7. *Conclusions*

Rejecting the idea of a scientific image clashing with a manifest one is rejecting the idea that we could change the categorial articulation of our conceptual structures. It is the conviction that while the rules of the game of giving and asking for reasons change, in the sense that we can modify our conceptual repertoires, they are still written in stone in the sense that we can't modify our categories. For as we have seen, if conceptual contents are constituted by norms, changing the categories of a conceptual structure means changing the structure of normativity. I hope I have offered enough reasons here to be wary of the idea that a similar recategorization would necessarily undermine our conceptual contents, even if they are construed as normatively articulated. I wish to suggest to the contrary that at least part of the task of philosophy may just consist in developing the expressive resources required to navigate such changes.

Giacomo Turbanti
Univeristy of Pisa
giacomo.turbanti@unipi.it

[28] As we have seen, this latter is part of what Sellars describes as acknowledging the picturing dimension of conceptual structures in chapter V of *Science and Metaphysics*.

References

Brandom, Robert, 2008, *Between Saying and Doing: Towards an Analytic Pragmatism*, Oxford University Press, Oxford.

Brandom, Robert, 2013, "Global Anti-Representationalism", Cambridge University Press, Cambridge.

Brandom, Robert, 2015, "Categories and Noumena: Two Kantian Axes of Sellars's Thought", in *From Empiricism to Expressivism*, Harvard University Press, Cambridge (MA): 30-98.

Carnap, Rudolph, 1950, *Logical Foundations of Probability*, Routledge & Kegan Paul, London.

deVries, Willem, 2005, *Wilfrid Sellars*, Routledge, London.

Eddington, Arthur Stanley, 1928, *The Nature of the Physical World*, The University Press, Cambridge.

Feynman, Richard, 1967, *The Character of Physical Law*, The MIT Press, Cambridge (MA). First published by the British Broadcasting Corporation, 1965.

Lewis, David Kellog, 1983, "New Work for a Theory of Universals", in *Australasian Journal of Philosophy*, 61: 343-377.

Millikan, Ruth Garrett, 2016, "Confessions of a Renegade Daughter", in O'Shea, James, ed., *Sellars and his Legacy*, Oxford University Press, Oxford: 117-129.

Quine, Willard V. Orman, 1960, *Word and Object*, The MIT Press, Cambridge (MA).

Rouse, Joseph, 2015, *Articulating the World: Conceptual Understanding and the Scientific Image*, University of Chicago Press, Chicago (IL).

Russell, Bertrand, 1913, "On the Notion of Cause", in *Proceedings of the Aristotelian Society*, 13: 1-26.

Sellars, Wilfrid, 1953, "Inference and Meaning", in *Mind*, 62, 247: 313-338.

Sellars, Wilfrid, 1954, "Some Reflections on Language Games", in *Philosophy of Science* 21, 3: 204-228.

Sellars, Wilfrid, 1957, "Counterfactuals, Dispositions, and the Causal Modalities", in *Minnesota Studies in the Philosophy of Science*, vol II, University of Minnesota Press, Minneapolis: 225-308.

Sellars, Wilfrid, 1961, "The Language of Theories", in Feigl, Herbert and G. Maxwell, Grover, eds., *Current Issues in The Philosophy of Science*, Holt, Rinehart, and Winston, New York (NY): 57-77.

Sellars, Wilfrid, 1962, "Philosophy and the Scientific Image of Man", in *Frontiers of Science and Philosophy*, University of Pittsburgh Press, Pittsburgh: 35-78. Reprinted in Sellars 1963: 1-40.

Sellars, Wilfrid, 1963, *Science, Perception and Reality*, Routledge and Kegan Paul, London.

Sellars, Wilfrid, 1964, "Induction as Vindication", in *Philosophy of Science*, 31, 3: 197-231.

Sellars, Wilfrid, 1968, *Science and Metaphysics: Variations on Kantian Themes*, Humanities Press, New York (NY).

Sellars, Wilfrid, 1969, "Metaphysics and the Concept of a Person", in K. Lambert ed., *The Logical Way of Doing Things*, Yale University Press, New Haven: 219-252.

Sellars, Wilfrid, 1970, "Towards a Theory of the Categories", in L. Foster and J.W. Swanson eds., *Experience and Theory*, University of Massachusetts Press, Amherst (MA): 55-78.

Sellars, Wilfrid, 1981, "Foundations for a Metaphysics of Pure Process" (The Carus Lectures), in *The Monist*, 64: 3-90.

Sider, Theodore, 2011, *Writing the Book of the World*, Oxford University Press, Oxford.

Tarski, Alfred, 1933, "Pojęcie prawdy w językach nauk dedukcyjnych", in *Prace Towarzystwa Naukowego Warszawskiego, Wydzial III Nauk Matematyczno-Fizycznych*, 34. Expanded English translation in Tarski 1983: 152-278.

Tarski, Alfred, 1944, "The Semantic Conception of Truth", in *Philosophy and Phenomenological Research* 4, 3: 13-47.

Tarski, Alfred, 1983, *Logic, Semantics, Metamathematics. Papers from 1923 to 1938*; Eng. tr. by J. Woodger, Hackett, Indianapolis (IN). Second edition edited and introduced by J. Corcoran.

van Fraassen, Bas, 1980, *The Scientific Image*, Claredon Press, Oxford.

Williamson, Timothy, 2007, *The Philosophy of Philosophy*, Blackwell, London.

Wittgenstein, Ludwig, 1922, *Tractatus Logico-Philosophicus*, Routledge & Kegan Paul, London.

Kant and Sellars on the unity of apperception

David Landy

Abstract: That Wilfrid Sellars claims that the framework of persons is not a descriptive framework, but a normative one is about as well known as any claim that he makes. This claim is at the core of the famous demand for a synoptic image that closes, "Philosophy and the Scientific Image of Man," makes its appearance at key moments in the grand argument of, "Empiricism and the Philosophy of Mind," and is the capstone of Sellars' engagement with Kant in *Science and Metaphysics*. Whereas mere things can be subject to ought-to-be rules – e.g. a clock ought to chime on the hour – to be a person, as Sellars understands it, is to be subject to ought-to-do rules – e.g. one ought to wind one's clocks to chime on the hour. *Prima facie*, though, there is more to being a person than just being subject to ought-to-do rules. For example, on at least some common ways of using 'person' to be a person is to have a unified consciousness, i.e. to be a single subject of a manifold of experience persisting through time. Arguably, that is what Kant takes a person to be. What I hope to show here is that it is what Sellars takes a person to be too. I.e. the exciting twist here is that as Sellars sees it being a single subject of experience persisting through time is being subject to a particular kind of ought-to-do rules, namely, those concepts-qua-inferential-rules that are the means by which we represent the world of causally-related objects existing in space and persisting through time.

Keywords: personal identity; practical reason; theoretical reason; transcendental deduction; synoptic image.

> Thus, in the union of pure speculative with pure practical reason in one cognition, the latter has primacy, assuming that this union is not *contingent* and discretionary but based a priori on reason itself and therefore *necessary*.
>
> (Kant 2015: 101)[1]

[1] My thanks to Samantha Matherne, Willem deVries, Tim Jankowiak, Micah Dubreuil, Luz Seiberth, Danilo Manca, and Giacomo Turbanti for reading and commenting on earlier versions of this paper. Thanks also to the participants of my Fall 2020 graduate seminar on Sellars at San Francisco State University for all of our discussions on this and related topics. Finally, thanks to the participants of the North American Kant Society Pacific Study Group 2021 meeting for their helpful questions and comments.

philinq X, 1-2022, pp. 73-96
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.393
Submitted: January 2021
Accepted: October 2021

That Wilfrid Sellars claims that the framework of persons is not a descriptive framework, but a normative one is about as well known as any claim that he makes. This claim is at the core of the famous demand for a synoptic image that closes "Philosophy and the Scientific Image of Man",[2] makes its appearance at key moments in the grand argument of "Empiricism and the Philosophy of Mind",[3] and is the capstone of Sellars' engagement with Kant in, *Science and Metaphysics*.[4] Whereas mere things can be subject to ought-to-be rules – e.g. a clock ought to chime on the hour – to be a person, as Sellars understands it, is to be subject to ought-to-do rules – e.g. one ought to wind one's clocks to chime on the hour.

Prima facie, though, there is more to being a person than *just* being subject to ought-to-do rules. For example, on at least some common ways of using 'person' to be a person is to have a unified consciousness, i.e. to be a single subject of a manifold of experience persisting through time. Arguably, that is what Kant takes a person to be.[5] What I hope to show here is that it is what Sellars takes a person to be too. I.e. the exciting twist here is that as Sellars sees it being a single subject of experience persisting through time *is* being subject to a particular kind of ought-to-do rules, namely, those concepts-qua-inferential-rules that are the means by which we represent the world of causally-related objects existing in space and persisting through time.

I take Sellars' reasons for holding this set of theses to be essentially Kantian, and so my procedure for explicating them will be to trace a single philosophical thread through both Kant's and Sellars' thinking surrounding these issues. I begin with the historical problematic to which Kant's Transcendental Deduction is intended as an answer. By what right does one apply the pure *a priori* concepts of the understanding? As the necessary means for representing the analytic unity of apperception, i.e. for representing oneself as the single subject of experience persisting through time. That leads to a consideration of the question of what the temporally-discursive experiences are of which one is supposed to be the single subject, and what the nature of the relation is of these experiences to such a subject. Here Sellars provides the answer. The question is ill formed. There is no *relation* of experiences to a subject because experiences are not themselves *things*. Rather, 'an experience' is a nominalization of the verb 'experiencing', which is itself a description of the act of a person. That thesis, then, brings us squarely to the question of what the framework

[2] Sellars 1963c: §114.

[3] Sellars 1963a: §12.

[4] Sellars 1967: Chapter VII, §1.

[5] E.g. A361. All citations from the *Critique* are to Kant 1998. I use the standard convention of citing the pages numbers of the original A and B editions.

of persons is, and why Sellars is so confident that it is an ineliminable feature of any future iteration of the synoptic image of the world. The answer to the latter question is that descriptive images themselves (scientific or manifest) are constituted by the rules that govern them, and it is only persons that can be subject to such rules.[6] So, in the end, we return to Kant's claim in the Transcendental Deduction that our representation of a world of causally-related objects existing in space and persisting through time is the means by which one represents oneself as the single subject of experience persisting through time. The descriptions that the scientific image provides are only possible, and necessary, because of what Kant would call their "ultimate principle": the framework of persons.

1. *The historical problematic*

To begin, consider the historical problematic surrounding the concept of the self with which Kant finds himself confronted.[7] Descartes seems to regard as valid an inference that moves from a premise of the form

(D1) [I think x] and [I think y] and [I think z],

to a conclusion of the form

(D2) [The I that thinks x] = [The I that thinks y] = [The I that thinks z].

That is, Descartes takes the fact that he can introspectively observe that he thinks x, and that he can introspectively observe that he thinks y, and that he can introspectively observe that he thinks z, to imply that it is one and the same thing, he, the thinking thing, that is the single subject of all of those thoughts.

> Is it not one and the same "I" who is now doubting almost everything, who nonetheless understands some things, denies everything else, desires to know more, is unwilling to be deceived, imagines many things even involuntarily, and is aware of thigs

[6] Willem DeVries points out in correspondence that there might also be more primitive representational systems that are likewise constituted by the rules that govern them, but which do not require persons to be subject to these rules. For example, the cries of monkeys that distinguish between predators above and predators below. What Sellars says about such systems in Sellars 1963b: §14-§16 is that while it is true that these particular cries are only explicable via appeal to the larger pattern of which they are a part, and that this is something that they share with the objects of ought-to-be rules, because ought-to-be rules are logically connected to ought-to-do rules, such cries are not genuinely rule governed. More below.

[7] This way of framing this problematic derives from Sellars' lectures on Kant at the University of Pittsburgh (Sellars 2002c) by way of Jay Rosenberg's lectures on Kant at the University of North Carolina (Rosenberg 2005). I have put it to use elsewhere as well (Landy 2015).

just as true as the fact that I exist, even if I am asleep all the time, and even if he who created me is doing all he can to deceive me? [. . .] The fact that it is I who am doubting and understanding and willing is so evident that I see no way of making it any clearer. (Descartes 1984: 19)

Of course, Hume finds the matter to be significantly less clear. He denies that this inference – from the introspective availability of certain experiences or perceptions to the identity of the subject of these experiences – is valid at all.

For my part, when I enter most intimately into what I call *myself*, I always stumble on some particular perception or other, of heat or cold, light or shade, love or hatred, pain or pleasure. I never catch *myself* at any time without a perception, and never can observe any thing but the perception. (T 1.4.6.3-4; SBN 252)[8]

What Hume points out here is that, when we introspect, we find exactly the matter that Descartes does – this or that perception – but that this is not sufficient to yield an experience of the self – something that endures through time and is the subject of these perceptions. Lacking such an experience of the self, Hume turns his attention to a different question. Given that I am able to introspect and find such-and-such a manifold of perceptions, what makes these perceptions, but not others distinctly *mine*? Hume's first answer to this question is,

> that the true idea of the human mind, is to consider it as a system of different perceptions or different existences, which are link'd together by the relation of cause and effect, and mutually produce, destroy, influence, and modify each other. (T 1.4.6.19; SBN 261)

The mind is a bundle of perceptions united by certain relations of cause and effect. In reconsidering his view in the appendix to the *Treatise*, Hume notoriously expresses his dissatisfaction with that account, but does not specify what the grounds of that dissatisfaction are.[9]

Kant, by contrast, is more than happy to express the grounds of *his* dissatisfaction with Hume's account. Most importantly for current purposes, Kant takes (D2) to be *analytic*, and dubs this representation of oneself the analytical unity of apperception. As Kant sees it, Hume is exactly right that merely adding together the manifold of representations represented in (D1) is not suffi-

[8] References to the *Treatise* are to Hume, *A Treatise of Human Nature*, ed. Norton and Norton, hereafter cited in the text as "T" followed by Book, part, section, and paragraph number, and to Hume, *A Treatise of Human Nature*, ed. Selby-Bigge, rev. by Nidditch, cited in the text as "SBN" followed by the page number.

[9] There are at least two dozen extant distinct interpretations of what bothers Hume in the Appendix.

cient for representing the subject of these representations as identical through time (D2). Where he goes wrong, however, is in inferring from this insight that such a subject cannot be represented at all.

> Namely, this thoroughgoing identity of the apperception of a manifold given in intuition contains a synthesis of the representations, and is possible only through the consciousness of this synthesis. For the empirical consciousness that accompanies different representations is by itself dispersed and without relation to the identity of the subject. The latter relation therefore does not yet come about by my accompanying each representation with consciousness, but rather by my adding one representation to the other and being conscious of their synthesis. Therefore it is only because I can combine a manifold of given representations in one consciousness that it is possible for me to represent the identity of the consciousness in these representations itself, i.e., the analytical unity of apperception is only possible under the presupposition of some synthetic one. (B133)

Kant holds that one can (and must) represent oneself as a single subject of experience persisting through time, and that the means for doing so is forming a single complex representation the components of which are the manifold of representations of which one can become introspectively aware.[10] That is, Kant holds that representing oneself as in (D2) is made possible by,

(K) I think [x + y +z].

Since [x + y +z] is a single unified representation, it is necessarily had by a single unified thinker. Since the components of [x + y + z] are the very representations that one finds via introspection, e.g. in (D1), (D1) and (D2) alike follow from (K): it is one and the same thinker that is the subject of x, y, and z. The representation of the self of the form presented in (K) is what Kant calls the synthetic or transcendental unity of apperception. We will return to this representation at the close of the current study.

In the meantime, notice that the success of Kant's strategy for resisting Hume's conclusion that one cannot so much as represent oneself as a single subject of experience persisting through time hangs on (among other things) Kant's claim that (D2) is analytic. It is analytic that the subjects of each of the representations that one finds in introspecting are all identical. Here is a famous passage to that effect.

[10] Strictly speaking, in the Transcendental Deduction, the unity of apperception under consideration is purely formal and so it does not yet include persistence through time. That condition is added to the concept of a person later in the *Critique*, once Kant has reintroduced our particular forms of intuition, Space and more fundamentally Time, in the Schematism.

> The I think must be able to accompany all my representations; for otherwise something would be represented in me that could not be thought at all, which is as much to say that the representation would either be impossible or else at least would be nothing for me. (B131-2)

By "The I think" here, Kant means a univocal representation, one that represents the same subject of experience as the subject of every representation that it accompanies. In light of Hume's rejection of the possibility of any such representation, however, it is worth asking why exactly Kant takes this thesis to be not only possible, but also true, and *analytic*. Hume appears to be able coherently to hold that there is no single subject of all of "my" representations persisting through time. If Kant is right, though, this appearance belies an underlying incoherence. What is that incoherence? Here, in answering this question, we will turn for the first time to Sellars.

2. *Sellars on sensations and thoughts*

To see why Sellars follows Kant in taking,

(D2) [The I that thinks x] = [The I that thinks y] = [The I that thinks z]

to be analytic, it will be helpful to begin again with Hume. Recall that Hume takes a distinctly *ontological* approach to accounting for the human mind. Hume holds that the mind is nothing other than the perceptions that compose it. Here is how he summarizes his own position in his anonymously-published abstract of the *Treatise*.

> He asserts, that the soul, as far as we can conceive it, is nothing but a system or train of different perceptions, those of heat and cold, love and anger, thoughts and sensations; all united together, but without any perfect simplicity or identity. *Des Cartes* maintained that thought was the essence of the mind; not this thought or that thought, but thought in general. This seems to be absolutely unintelligible, since every thing, that exists, is particular: And therefore it must be our several particular perceptions, that compose the mind. I say, *compose* the mind, not *belong* to it. The mind is not a substance, in which the perceptions inhere. (A 28; SBN 657-658)

Hume's portrayal of Descartes here is almost certainly inaccurate, but putting that matter aside, notice again that Hume holds that the mind is a complex composed entirely of its perceptions. I.e. he holds that our primary understanding of perceptions categorizes them as *things* and the question of the unity of the subject of experience hinges on the further question of what

the *relation* of these things is to that subject. In the end, Hume denies that there can be any such relation, and so settles (temporarily, at least) on his view that the mind just is a bundle of perceptions. I.e. the mind is *composed* of its perceptions, which stand in certain relations only to each other.

What is significant about this way of understanding Hume's account is that it rests on what Sellars takes to be a false premise: that our primary understanding of mental representations categorizes them as *things*. In place of such an understanding, Sellars urges that we interpret the idioms of mental representa*tions* as nominalizations of verbs of represent*ing*, which apply first-and-foremost to *persons*.

> [I]t is surely implausible to take such statements as
>
> Tom has a feeling
>
> to be anything but a derivative (but legitimate) way of saying what is said adequately and non-relationally by such statements as
>
> Tom feels…
>
> Thus, in general,
>
> Tom has a V-tion,
>
> where 'V-tion' is a verbal noun for a kind of "experience", would be a derivative (but legitimate) way of saying what is said adequately and non-relationally by
>
> Tom Vs. (Sellars 2002a: 313-314)

As Sellars sees it, the vocabulary of ideas, perceptions, etc., while it *appears* to categorize mental representations as things, is actually a derivative idiom. Primarily, thinking, perceiving, representing, etc. are *acts* of *persons*. Just as we might say,

> Dave wore a smirk

and thereby invite the question of what the relation is between Dave and the smirk that he wore, such a question is easily avoided by noticing that 'a smirk' is a nominalization of the verb 'smirking'. Smirks do not exist without the persons that wear them, not because the existence of smirks depends in some metaphysical way on the existence of persons, but rather because 'a smirk' is a derivative way of representing the act of smirking, which is something that persons do. The above is more perspicuously put as,

> Dave smirked.

Similarly, Sellars suggests that 'I think x' is not a statement of a relation between me and *the thought* x, but rather a representation of me as think*ing*. More specifically, it is the representation of me as thinking *x*, where 'x' is also not a *thing* (say, a content) to which my thought relates, but is rather a functional classification of my thinking.

> The thought that snow is white occurred to Jones,

which is doubly relational in appearance, turns out to have as its foundation the non-relation state of affairs expressed by

> Jones ·snow is white·ed. (Sellars 2002a: 318)

Just as single quotes mention the word between them qua a word in a particular language – e.g. 'dog' has three letters – Sellars' dot quotes individuate words according to their inferential role – e.g. German 'rot's are ·red·s.[11] I.e. 'rot' plays the same inferential role in German as 'red' does in English. Sellars' first point here is that in attributing the thought 'snow is white' to Jones, firstly, we should conceive of that thought not as an entity distinct from Jones himself, but rather an act of thinking that Jones performs. His second point is that in in specifying that Jones is thinking that snow is white, we are classifying that act of thinking as playing a certain inferential role: it is a thinking of a snow-is-white kind, or is a snow-is-white thinking.

To return to our analogy, rather than interpret,

> Dave wore a smirk that was wry,

as a relation between Dave, the smirk that he wore, and its wryness, we should understand it as a derivative way of expressing an adverbial classification more perspicuously represented by,

> Dave smirked wryly.

Dave is the only *thing* represented in that proposition; he is represented as having acted, he smirked; and he acted in a particular way, he smirked wryly.

Analogously, then,

(D1) [I think x] and [I think y] and [I think z],

[11] I use "inferential role" here as shorthand for the entire complex of language-entry, language-language, and language-exit transitions that Sellars takes to constitute a language. It is important to note, however, that this shorthand belies the fact, about which Sellars is explicit and emphatic, that language-entry and language-exit moves are not inferences.

which attributes a manifold of representations to me is not the statement of a relation between a thing, me, to some other things, the thoughts, x, y, and z. Rather, it is a functional classification of the various acts of a person, me, as an act of functional kind x, an act of functional kind y, and an act of functional kind z. The important point here is that the notion of a person is logically prior to the notion of that person's thoughts in just the same way that the notion of a person is logically prior to the notion of a smirk. 'A thought,' just like 'a smirk' is a nominalization of its corresponding verb, 'thinking', and no account is needed of the relation of a person to their thoughts.

Of course, Sellars famously insists on a distinction between thoughts (conceptual representations the content of which can be described functionally) and sense impressions (which do not have a content *per se*, but which are characterized firstly according to their causal role, and then also by their intrinsic characteristics). What I have been concerned with above is thoughts, but Sellars offers an analogous treatment of sense impressions.

In this perspicuous language we would not say,

> Tom senses a red triangle

but

> Tom a-red-triangles

where the verb 'a-red-triangle' stands for that kind of sensing which is brought about in standard conditions, and in standard perceivers, by the presence of a literally red and triangular object. (Sellars 2002a: 317)

To say that a subject has a certain sensation is not to assert that a relation holds between that subject and that sensation, but is rather to classify an act of that person's as being of the kind that is typical of persons in such-and-such circumstances. In the case of both thoughts and sensations, the idiom of mental *representations* is derivative of the idiom of *persons* and their acts of mental *representing*, and is used primarily to classify such acts, either according to their inferential role in the case of thoughts, or according to their causal role in the case of sensations. So, our concepts of persons are logically prior to our concepts of their representations, and the question of the relation of former to the latter ought not to arise. Thinking x and thinking y (or sensing x and sensing y) are ways that *I* act just as smirking or winking are ways that I can arrange my facial features. My thought x and my thought y depend on me in the same way that my smirk or my wink do.

Of course, this interpretation of the deep logical grammar of representation is not by itself sufficient for accounting for the analyticity of the analytical

unity of apperception. The question of whether the person who winked was the same as the person that smirked can be a perfectly coherent one in certain circumstances. Analogously, the question of whether the person who thought "Snow is white" was the same person as that who thought "Grass is green" can be a perfectly coherent one. There is more work to be done in telling Kant's and Sellars' story. For the moment, however, another task has it pushed its way to the top of our agenda.

Readers familiar with Sellars might note that the account just offered of the idiom of representation, both that of sensation and of conceptual representation, as being derivative of the idiom of persons and their acts of representing, is one that has thus far concerned only the manifest image. That is, this account has cast representing as something that persons do, and 'persons' is very much a manifest-image category. One might wonder, then, what happens to this account when the descriptive component of the manifest image is replaced by that of the scientific image. Does the category of persons persist? Is it guaranteed to persist across all such changes in images?

3. *The persistence of 'persons'*

In answering these questions, it is important to recall that the manifest image is itself a synoptic image: it has both a descriptive component and a normative one. I.e. one aspect of the manifest image is the picture of the world that it contains. This picture represents the world as being composed of tables and chairs, elephants and mice, etc. Another aspect of it is non-descriptive, and consists of the vocabulary of reasons, entitlements, commitments, etc. Demands for explanation and prediction (more on the source of which soon) reveal the inadequacy of the *descriptive* component of the manifest image, and thus the scientific image is created as its replacement, but there are no such parallel pressures put on its normative aspect. While certain philosophical considerations might put explanatory pressures on our account of norms – e.g. Sellars take very seriously the need to account for the motivational power of normative beliefs – the pressure to produce a picture of the world of increasing accuracy and predictive and explanatory success does not, as far as Sellars is concerned, put a corresponding pressure on the normative idiom of the manifest image. That idiom is never intended as descriptive, and so the need for an improved description of the world is largely orthogonal to its functioning.[12]

[12] Sellars famously comments on this confusion of these two aspects of the manifest image: "Now the idea that epistemic facts can be analysed without remainder – even 'in principle' – into non-epistemic facts […] is, I believe, a radical mistake – a mistake of a piece with the so-called 'naturalistic fallacy' in ethics." (Sellars 1963a: §12)

This difference between the descriptive aspect of the manifest image and its normative aspect is what requires that we pay careful attention to the distinction between sensory representing and conceptual representing. As we noted above, to classify a mental representing as a sensing, and as a sensing of a particular kind, is to give (or issue a promissory note for giving) an intrinsic characterization of it. In "Tom a-red-triangles" the verb 'a-red-triangle' stands for that kind of sensing which is brought about in standard conditions, and in standard perceivers, by the presence of a literally red and triangular object. As such, sensings are themselves "mere" states of the experiencing subject, and their ultimate representation is the province of the scientific image.[13] By contrast, in "Jones ·snow is white·ed" the verb '·snow is white·ed' classifies Jones' thought as playing a certain inferential role, or as being subject to certain *rules* of inference. For example, it takes Jones to be subject to criticism if he also thinks that snow is blue, or if he does not also think that snow is the same color as clouds, etc.

According to the manifest image, it is persons that are the logical subjects of ascriptions of mental states, both sensory mental states and conceptual ones. In the case of sensory states, the scientific image replaces the manifest image, and so the logical subjects of the scientific image replace the logical subjects of the manifest image. Persons qua the logical subjects of sensory states become persons qua collections of atoms in the void, and as Sellars sees it, eventually persons qua absolute processes. In the case of conceptual representings or thoughts or thinkings, however, since these are not *per se* ontological, but rather normative, their logical subject persists through changes in our conceptions of its material constitution. It is *a person* that is subject to rules of performance should they fail to satisfy their inferential commitments, and our concept of a person is not the concept of a particular kind of matter-of-factually describable substance, but rather is the concept of that which is subject to such norms. To use a version of one of Sellars' preferred analogies, the rules that govern playable moves in a game of chess apply equally to human, computer, and alien players alike, regardless of whether they are made primarily of carbon, silicon, or more exotic kinds of matter. What makes one a player of chess, is that one is subject to the rules of chess. Analogously, what makes one a thinking person, is that one is subject to the (linguistic and/or conceptual) inferential rules that constitute the content of our thoughts.

For as was pointed out, not only are concepts pertaining to conceptual representations analogical counterparts of concepts pertaining to verbal be-

[13] Thereby hangs a tale. I have attempted to relate that epic yarn in Landy (2019). Others have told their own version of it as well, e.g. Rosenberg (1982), Rosenthal (2015), Seibt (2015), and Seibt (2000).

haviour but, which is more important, the latter concepts are concerned with correctnesses and uniformities of linguistic configurations, extra-linguistic objects and non-linguistic behaviour. As for the "qualitative content" of these configurations, it must, we have said, be such as to be capable of taking part in these configurations. (Sellars 1967: 173)

To classify a mental state as a conceptual representing is to classify it according to what makes it correct or incorrect, what one is obliged or forbidden to represent in virtue of endorsing such a representing, what warrants such representing and what one is required to do in virtue of having so represented. As Sellars notes here, such classifications are not descriptions, and entail descriptions of the world only insofar as they require that the world be such as to make possible such classifications. To return to the chess analogy, classification of some worldly object as a chess piece leaves enough latitude for the classification to apply to such diverse objects as pieces of wood, luxury automobiles, pixels on a screen, and human beings. Thus, it seems as though while the replacement of the descriptive component of the manifest image might entail a recategorization of our concepts of acts of *sensing* as distributed across logically complex subjects, there is no reason to think that such a recategorization will be required for our *conceptual* representings.

Sellars, however, argues for an even stronger conclusion. Not only do we have no reason to think that conceptual representings and the normative structure of persons of which they are a part will require replacement, but we can also know that they will not. I.e. the concept of persons is guaranteed to persist through all subsequent replacements of descriptive content by the scientific image.[14]

[14] In response to a question about whether 'person' would be a category in Sellars' "ultimate ontology" following one of his lectures at Notre Dame, Sellars proclaims:: "No, I mean it will. On the contrary. Ontology is functioning, again, in an ambiguous way here. Of course there are persons. The question is: could a person have the kind of features that in practical reasoning we essentially conceive of them as having and still, in some sense, be pluralities? That's why, in my essay, 'Towards a Metaphysics of the Person,' I discussed Kant's Paralogisms. In the paper, 'This I or He or It that thinks…' the presidential address, at the end of it I, again go into Kant's Paralogisms. Because couldn't a person have exactly those features which are required by a normative view and still be a plurality, be a system, as opposed to a Cartesian simple? See, this is the old question going back to Plato of the simplicity of the soul, roughly. And that's what Kant was arguing about in the Paralogisms. And that's what I was arguing about. Kant, in fact, said that a person *could* be a system and still have those features which his ethics required him to have, and that's exactly what I say, and that's the only sense in which persons would not be ultimate simples. If you mean by 'what your ontology contains' the ultimate simples that it contains, you see, then persons might not be ultimate simples. But, as I said, that's the old classical issue of the simplicity of the soul" (Sellars 2018: 300). Notice that Sellars' response begins by pointing out that "ontology" is functioning in an ambiguous way in the question. This is because he does not take 'person' to be an ontological, or descriptive, category at all. Thus, whatever changes we make to the descriptive component of our image of the world, the category of 'persons' remains in tact, and as we are about to see, must remain in tact.

But this seems to leave the door wide open, for, as was pointed out, almost anything can be used to play the game of chess. Thus we might be inclined to say that almost anything could be the material cause or "matter" (in the Aristotelian sense) of the configurations which are conceptual representations. Yet things are not quite so simple. For though we have been emphasizing that the candid thinkings-out-loud which are the models for mental acts are not *actions*, and that the mental acts for which they are the models are not *actions* but rather acts in the Aristotelian sense, nevertheless, though we have not been emphasizing the point, there *are* mental actions. And, indeed, if there were no *actions* pertaining to thinking […] there would be no thinking […] but at best processes which, however sophisticated would be simulations of thought. (Sellars 1967: 173-174)

For all that we have said about conceptual acts of representing to this point, such acts could be governed entirely by what Sellars calls ought-to-be rules. Ought-to-be rules are rules that are applied without the object of those rules having to be in any sense aware of those rules. E.g. a clock ought to be wound so that it chimes on the hour. That rule does not require any *action* from the clock, even if we describe the chiming of a clock as an act in Sellars' Aristotelian sense. Sellars' point here, however, is that the clock's being subject to that ought-to-be rule does require something of someone. Whoever is responsible for the winding of the clock is subject to a corresponding ought-to-*do* rule: he or she ought to wind the clock so that it chimes on the hour.

Now ought-to-be's (or rules of criticism as I shall also call them), though categorical in form, point beyond themselves in two ways. In the first place they imply (in some sense of this protean term) a reason, a because clause. The exploration of this theme would seem to take us back to the excluded topic of hypothetical imperatives. In the second place, though ought-to-be's are carefully to be distinguished from ought-to-do's they have an essential connection with them. The connection is, roughly, that ought-to-be's imply ought-to-do's. Thus the ought-to-be about clock chimes implies, roughly,

> (Other things being equal and where possible) one ought to bring it about that clock chimes strike on the quarter hour.

This rule belongs in our previous category, and is a rule of action. As such it requires that the item to which it applies (persons rather than chimes) have the appropriate concepts or recognitional capacities. (Sellars 1969: 508)

In contrast to the ought-to-be rule that governs the clock, the ought-to-do rule that governs the person responsible for winding the clock does require that the person that is its subject is able, in some suitable sense, to represent that rule as governing his or her behavior. Paradigmatically, the subject of an

ought-to-do rule is able to represent their actions as being governed by rules of practical reasoning. For example, they can employ the classic form of a practical syllogism.

1. I shall bring about E.
2. Bringing about E implies doing A.
3. Therefore, I shall do A.[15]

In the example of the person responsible for winding the clocks, such a syllogism might be constructed as follows.

1'. I, a clock winder, shall bring it about that clocks chime on the hour.
2'. Bringing it about that clocks chime on the hour implies winding the clocks daily.
3'. I shall wind the clocks daily.

Ceteris paribus, forming the intention expressed in 3' results in my subsequently forming what Sellars calls a here-and-now volition to wind a clock, which in turn results in my then and there winding a clock. In the case of a conceptual agent constructing a descriptive image of the world, as we saw Sellars point out above, the inferences in question, "are concerned with correctnesses and uniformities of linguistic configurations, extra-linguistic objects and non-linguistic behaviour." An example of a practical syllogism with these concerns would be something like the following.

1''. I, a subject of experience, shall bring it about that my thinking is consistent.
2''. Bringing it about that my thoughts are consistent implies rescinding either my belief that it has rained (because the streets are wet) or my belief that it has not rained (because I have not heard rain in some time).
3''. Therefore, I shall rescind either my belief that it has rained (because the streets are wet) or my belief that it has not rained (because I have not hear rain in some time).

Ceteris paribus, forming the intention expressed in 3'' results in my subsequently forming a here-and-now volition to rescind one of these beliefs (perhaps after checking the weather, seeing if the neighbor has been running his sprinkler, etc.), which in turn results in my then and there rescinding one of those beliefs.

Notice that each of the syllogisms that we have considered concludes with

[15] Cf. Sellars 2002b: 308.

the formation of an intention (indicated by the 'shall' that appears in each of them), but does not reach its fruition until that intention itself results in a volition, which in turn results in the action intended. While this terminus is in no sense guaranteed by the mere fact of the reasoning having occurred, it is part of the logic of practical reasoning, that such actions are its *ceteris-paribus* results. I.e. one important difference between theoretical and practical reasoning is that practical reasoning aims at, and *ceteris paribus* results in, action. As Sellars sees it, that is because practical reason begins and ends with intentions, and intentions would not be intentions if they did not bear this relation to volition and in turn to actions. This connection of practical reasoning to action brings us back to the connection between the subject of experience and practical reasoning.

Recall that the dialectic we have been tracing ran thusly. We noticed that Sellars holds that the logical subjects of *thoughts*, persons, are guaranteed to persist through all replacements of the descriptive content of the manifest image by the progressively more explanatory scientific image. We wondered what the grounds for this claim were, and began investigating them by noticing that Sellars' takes it to be an essential feature of conceptual thinking that it is subject to certain norms, rules of inference. This, in turn, led us to distinguish ought-to-be-rules from ought-to-do rules, and to further note that something's being subject to the former implies that something, a *person*, is also subject to the latter. The outlines of our guarantee are now coming into focus. Conceptual thinking is rule-governed representation; it is persons that are the subject of such rules; thus conceptual thinking guarantees the persistence of persons.[16]

The question remains, however, whether anything here further guarantees the unity of *a person across a variety of representations* (especially through time), which is the essential feature the analytic unity of apperception. Sellars' answer is that it does. Specifically, it is the essentially first-personal role of *intentions* in the pieces of practical reasoning that constitute conceptual thinking that does so.

> These considerations highlight the fact that the intention expressed by a 'shall' statement is invariably the speaker's intention. Thus,
>
> Tom shall do A
>
> Expresses the speaker's intention that Tom do A. This 'first person' feature of intentions consists in part in their relation to the
>
> I shall do

[16] In fact, it guarantees the persistence not just of persons but also the *concept* of persons because, as we are about to see, a person just is that which conceives of itself using the concept person.

Which can become the commitment to do something *here* and *now* which is volition. (Sellars 1967: 184-185)

It is part of the logic of practical reasoning that the subject of the intention that serves as a premise must be the same as the subject of the intention that is the conclusion (and that eventually, *ceteris paribus*, performs the action so intended). For example, such arguments would not be valid if it were one person that intended to achieve some end, but another that took up the means to that end. E.g. the following inference is obviously invalid.

1. Dave shall achieve end E.
2. M is the means to E.
3. Therefore, Sherlock shall pursue M.

The very idea of practical reasoning depends on the univocality of the subject of such reasoning, that it is the very same person that adopts both the ends and means at issue. Furthermore, though, this form of reasoning supposes not just that it is *some* one person that pursues both these ends and means, but that it is some *I* that does so. Notice the difference between the following two arguments.

1. Dave has end E.	1'. I shall achieve end E.
2. M is the means to E.	2'. M is the means to E.
3. Dave pursues M.	3'. I shall pursue M.

1 and 3 are third-personal reports on Dave's intention and do not themselves result in the formation of any new intentions or volitions, and do not directly lead to any actions. 1' and 3', however, are each the expression of an intention, and 3' *is* importantly a new intention that is formed in the course of this reasoning, and does, *ceteris paribus*, lead to the appropriate volition and action.[17] So, effective practical reasoning presupposes the ability to represent *myself* as the *single* subject of practical norms.

Combining this conclusion with our earlier observation that Sellars understands conceptual representation as representation according to conceptual norms, we arrive at the conclusion that for there to be any conceptual repre-

[17] Notice that if in the following argument, 1'' and 3'' *are* intentions, but they are the speaker's intentions to shape Dave's behavior.

Dave shall achieve end E.
M is the means to E.
Dave shall pursue M.

I.e. the intentions expressed here are the speaker's intentions *that Dave achieve end E* and *that Dave pursue M*. My thanks to Willem DeVries for his guidance in sorting out these three arguments.

sentation at all presupposes that those representations are the representations of a single subject of experience persisting through time, i.e. the "existence" of persons, i.e. the analytic unity of apperception. If I am subject to ought-to-do rules because I subject my actions to practical reasoning, then it must be one and the same 'I' that is the subject of the intentions that are the premises of those syllogisms as it is the subject of the intentions that are their conclusions. If, however, it is *by* subjecting my actions to such bits of reasoning that I represent myself as a person at all, then it does turn out to be analytic that the I that is subject of any one of my representations must be identical to the I that is the subject of any other. That is, if what it is to be the subject of a conceptual representing is to be the subject of a piece of practical reasoning concerning that representing, and all practical reasoning requires that the subject of each of its steps is one and the same person, then being the subject of a conceptual representing requires a unity of the subject of piece of practical reasoning through all of its steps.

Of course, if we are to conceive of persons as being the subjects of practical means-end reasoning, and of their conceptual representing as actions subject to criticism in light of the ought-to-do rules adopted via such reasoning, then we must also ask what the *ends* being pursued in such reasoning are. If the unity of the conceptualizing subject is the unity of a person subject to norms of practical reasoning, then this raises the question of what the *ends* are of such a reasoner.[18] To what end does one engage in conceptual representation at all?

4. *The end of theoretical reasoning*

Before we can delve into Kant's and Sellars' answer to this question, it will be helpful to say a little bit more about the kind of representation that they both take conceptual representation to be. Both Kant and Sellars begin their treatments of concepts by noting the different roles played by that which serves as the cognitive analogue of the subjects and predicates of linguistic propositions. Here is Kant.

> [T]he cognition of every, at least human, understanding is a cognition through concepts, not intuitive but discursive. All intuitions, as sensible, rest on affections, concepts therefore on functions. By a function, however, I understand the unity of the action of ordering different representations under a common one. Concepts are therefore grounded on the spontaneity of thinking, as sensible intuitions are grounded on

[18] Certainly, there is more than just one end of all practical reasoning, but what we are pursuing here is what Kant would call its supreme principle.

the receptivity of impressions. Now the understanding can make no other use of these concepts than that of judging by means of them. Since no representation pertains to the object immediately except intuition alone, a concept is thus never immediately related to an object, but is always related to some other representation of it (whether that be an intuition or itself already a concept). Judgment is therefore the mediate cognition of an object, hence the representation of a representation of it. In every judgment, there is a concept that holds of many, and that among this many also comprehends a given representation, which is then related immediately to the object. (A68/B93)

Whereas an intuition is a determinate singular representation that pertains immediately to the object (roughly, refers to it), a concept is a kind of meta-representation. It serves as a function that takes certain intuitions as its inputs and outputs other intuitions. More specifically, as I have argued elsewhere, Kant takes concepts to be inferential rules (Landy 2015). Intuitions related to one another via such rules collectively form a picture of the world of objects existing in space and persisting through time as necessarily connected to each other. To understand how this account of mental representation is supposed to work, consider the following inference:

1. *x* is to the north of *y*.

---

2. *y* is to the south of *x*.

At first blush, this inference appears to be an enthymeme. It is not an example of *modus ponens*, *modus tollens*, conjunction elimination, or any other formal rule of inference. Since it is not valid in virtue of its logical form, the thinking goes, the ring of validity that it has can only be due to a suppressed premise. The valid argument that this one stands in for is really the following:

3. *x* is to the north of *y*.
4. If *x* is to the north of *y*, then *y* is to the south of *x*.

---

5. *y* is to the south of *x*.

The problem now is that while 4 does make for a formally valid inference when paired with 3 and 5, there is a closely related inference for which 4 is of no help.

6. Suppose *x were* to the north of *y*.

---

7. Then, *y would be* to the south of *x*.

4 is of no help here because it concerns only the actual relation of *x* to *y*. It concerns only what is the case if *x* actually is to the north of *y*, not what would be the case were *x* to be north of *y*. In order to validate the inference from 6 to 7, what is needed is a proposition that applies to these counterfactual situations as well. To this end, one might be tempted to offer,

8. If *x* is to the north of *y*, then *necessarily*, *y* is to the south of *x*.

Notice, however, that if this premise is accepted, so can the meta-level rule of inference,

9. '*x* is to the north of *y*' implies '*y* is to the south of *x*'.

That is, if 8 is true, then 1 can never be true where 2 is false. Thus, the inference from 1 to 2 is valid. Thus, in any system of representation robust enough to encompass the truth of some subjunctive conditionals, there will be principles of inference corresponding to these conditionals that are valid in virtue of something other than merely their logical form. Kant's proposal as I understand it is that it is *by* licensing such inferences, in the form of deploying concepts-as-inferential-rules, that we represent spatiotemporal objects as necessarily connected to one another. In the case above, it is by licensing the inference from 1 to 2 that we represent x and y as standing in a certain lawful spatial relation. Sellars offers an example of using a similar logical mechanism to represent the world's causal structure.

> Law-like statements, therefore, are empirically based principles of inference which authorize such inferences as, to use a crude example, 'Lightning now, therefore thunder shortly.' It also authorizes such conditionals as 'If there had been lightning then, there would have been thunder shortly' and such statements as 'There was thunder then because there had been lightning shortly before' and 'That there was lightning shortly before made it necessary that there be thunder then. (Sellars 2002b: 313)

Consider again the inferences at issue.

1'. There was just a flash of lightning.

2'. There will be a clap of thunder soon.

On its face, this argument is an enthymeme, in need of a supporting premise such as:

3'. If there was just a flash of lightning, there will be a clap of thunder soon.

Of course, while 3' might be sufficient to complete this argument, the closely related argument,

4'. Suppose there had been a flash of lightning just now.

5'. Then there would be a clap of thunder soon.

requires instead,

6'. If there were a flash of lightning, then, there would be a clap of thunder,

which implies that meta-level rule of inference,

7'. 'There was just a flash of lightning' implies "There will be a clap of thunder soon".

Thus, Sellars concludes with Kant that the way that we represent the world of spatiotemporal objects necessarily connected to one another by causal laws is by relating representations of these objects to one another (intuitions in Kant's case, names in Sellars') via inferential rules. To put it in the idiom of Wittgenstein's *Tractatus*, that we relate "*x*" and "*y*" to each other via material-inferential rules pictures that *x* and *y* are necessarily connected to each other.[19]

With that said, the question that we posed at the close of the previous section can now be put into a more specific form: what is the end for which representing the world of necessarily-connected spatiotemporal objects as such is the means? Kant's answer to this question brings us full circle. Recall that I began this study by noticing Kant follow Hume in rejecting the inference they both find in Descartes from a premise of the form,

(D1) [I think x] and [I think y] and [I think z]

to a conclusion of the form,

(D2) [The I that thinks x] = [The I that thinks y] = [The I that thinks z].

We further noted at the time that while Kant rejects this *inference*, he

[19] Nevertheless, as we shall see, the full flavor of actual modal discourse involves the way in which sentences in the first level language game containing modal words parallel sentences containing rule words ('may', 'ought', 'permitted', etc.) in the syntactical metalanguage. This parallelism is quite intelligible once one notes that the moves which are signalized in the object language by sentences containing modal words, are enjoined (permitted, etc.) by sentences containing rule words in the syntactical metalanguage (Sellars 1963b: 209).

nonetheless takes (D2) to be analytic, and so sets out to discover how it is that we can represent ourselves as single subjects of experience persisting through time (D2, the analytic unity of apperception), if not by merely representing the manifold of perceptions that constitute such a subject (D1). What Kant discovers is,

(K) I think [x + y +z].

Since [x + y +z] is a single unified representation, it is necessarily had by a single unified thinker. Since the components of [x + y + z] are the very representations that one finds via introspection, e.g. in (D1), (D2) follows: it is one and the same thinker that is the subject of x, y, and z. The representation of the self of the form presented in (K) is what Kant calls the synthetic or transcendental unity of apperception. Kant reports his result as being the surprising claim that the analytic unity of apperception depends on the synthetic one. There is another surprising claim implicit here as well, though. The representation [x + y + z] is the representation of *an object*, which as Kant reveals later is itself an instance of the representation of *the world*, i.e. the necessary connection of all spatiotemporal objects. So, what is required for representing oneself as the single subject of a manifold of representations (the analytic unity of apperception) is to represent oneself as the single subject of a single complex representation (the synthetic unity of apperception) of the world of necessarily-connected spatiotemporal objects.

> The supreme principle of all intuition in relation to the understanding is that all the manifold of intuition stand under conditions of the original synthetic unity of apperception. All the manifold representations of intuition stand under [this principle] insofar as they must be capable of being combined in one consciousness; for without that nothing could be thought or cognized through them, since the given representations would not have in common the act of apperception, I think, and thereby would not be grasped together in a self consciousness.
>
> Understanding is, generally speaking, the faculty of cognitions. These consist in the determinate relation of given representations to an object. An object, however, is that in the concept of which the manifold of a given intuition is united. (B136-B137)

What Kant is saying here is that the justification (the answer to his famous *quid juris*) for representing the necessarily-connected world of spatiotemporal objects is that it makes possible the analytic unity of apperception! That is, the end of theoretical reasoning, as Kant understands it, is the representation of oneself as the single subject of experience persisting through time. Since we have seen that and why Kant and Sellars take that representation to be

analytic, we can now also see why Kant takes the theoretical reasoning to be a non-optional form of our cognition: it is the necessary means to the analytic end of representing ourselves.

The only question remaining in this study, then, is whether Sellars agrees. I have suggested that Sellars agrees with Kant about the nature of the analytic unity of apperception. I have also suggested that he agrees with Kant about the nature of theoretical reasoning. Does he also draw the justificatory connection between these two pieces of representation that Kant does? The first thing to note in answering this question is that there may not be just one answer to it. For example, James O'Shea has argued that Sellars' position on this issue changed over the course of his career (2007: 129-136). As O'Shea sees it, Sellars gives one answer to the question of what justifies theoretical reasoning in his early essays including "Empiricism and the Philosophy of Mind" (1956), and a different one in later essays such as "On Accepting First Principles" (1988). For current purposes, I will focus on the view that O'Shea takes Sellars to articulate only in the later essays. To that end, here is Sellars in 1975 articulating how he would respond to a demand to justify employing the kind of cognitive framework that we have just been discussing.

> It must, indeed, be granted that principles pertaining to the epistemic authority of perceptual and memory beliefs are not the sort of thing which *could* be arrived at by inductive reasoning from perceptual belief. But the best way to make this point is positive. *We have to be in this framework to be thinking and perceiving beings at all.* (Sellars 1975: §45)

Any particular set of concepts-qua-material-inferential rules will be adopted on the grounds that they best explain observed empirical generalizations. But what justifies our inductive and explanatory practices themselves? Sellars' answer is that what justifies the use of this framework is that it is necessary for *being a thinking and perceiving being at all.* He elaborates.

> I pointed out a moment ago that we have to be in the framework of these (and other) principles to be thinking, perceiving, and now I add, acting beings at all. But surely this makes it clear that the exploration of these principles is but part and parcel of the task of explicating the concept of a rational animal or, in *VB* terms, of a language-using organism whose language is *about* the world in which it is *used*. (Sellars 1975: §46)

Notice that Sellars makes special mention of the fact that employing a conceptual framework is a necessary condition not only for thinking and perceiving, but also for acting. That addition fits with the general outline of the dialectic we have been sketching. The unity of the experiencing subject is the unity

of the subject of inferential ought-to-do rules, it is the unity of a conceptual *agent*. What we can now see is the point that Sellars makes in the second sentence here: that the unity of such conceptual agents is achieved via representing oneself as the single subject of experience *of a causally-structured world* (of objects existing in space and persisting through time). So, indeed, Sellars does share Kant's views not just of the experiencing subject, and of the experienced world, but also of the justificatory connection between these.

David Landy
San Francisco State University
landy@sfsu.edu

References

Descartes, Rene, 1984, *The Philosophical Writings of Descartes*, vol. 2, Eng. transl. by J. Cottingham, R. Stoothoff and D. Murdoch, Cambridge University Press, Cambridge.

Hume, David, 1974, *A Treatise of Human Nature*, ed. by L.A. Selby-Bigge, Oxford University Press, Oxford.

Hume, David, 2000, *A Treatise of Human Nature*, ed. by D. Fate Norton and M.J. Norton, Oxford University Press, New York.

Kant, Immanuel, 2015, *Critique of Practical Reason*, Eng. transl. by Mary Gregor, Cambridge University Press, Cambridge.

Kant, Immanuel, 1998, *Critique of Pure Reason*, Eng. transl. by Paul Guyer and Allen W. Wood, Cambridge University Press, Cambridge.

Landy, David, 2019, "Sellars' Argument for an Ontology of Absolute Processes", in *Journal for the History of Analytical Philosophy*, 7, 1: 1-25.

Landy, David, 2017, *Hume's Science of Human Nature*, Routledge, New York.

Landy, David, 2015, *Kant's Inferentialism*, Routledge, New York.

O'Shea, James, 2007, *Wilfrid Sellars: Naturalism with a Normative Turn*, Polity Press, Cambridge.

Rosenberg, Jay F., 1982, "The Place of Color in the Scheme of Things: A Roadmap to Sellars' Carus Lectures", in *The Monist*, 65: 315-335.

Rosenthal, David, 2015, "Quality Spaces, Relocation, and Grain," in James O'Shea, ed. *Sellars and His Legacy*, Oxford University Press, Oxford: 149-185.

Seibt, Johanna, 2015, "How to Naturalize Sensory Consciousness and Intentionality within a Process Monism with Normativity Gradient: A Reading of Sellars", in James O'Shea, ed. *Sellars and His Legacy*, Oxford University Press, Oxford: 186-223.

Seibt, Johanna, 2000, "Pure Processes and Projective Metaphysics", in *Philosophical Studies*, 101: 253-289.

Sellars, Wilfrid, 2018, *Wilfrid Sellars Notre Dame Lectures 1969-1986*, ed. by P.Amaral and D. Landy, Ridgeview Publishing Co., Atascadero.

Sellars, Wilfrid, 2002a, "Metaphysics and the Concept of a Person", in Jeffrey F. Sicha, ed. *Kant's Transcendental Metaphysics*, Ridgeview Publishing, Atascadero: 304-329.

Sellars, Wilfrid, 2002b, "On Accepting First Principles", in Jeffrey F. Sicha, ed. *Kant's Transcendental Metaphysics*, Ridgeview Publishing, Atascadero: 461-571.

Sellars, Wilfrid, 1975, "The Structure of Knowledge", in Hector-Neri Castañeda, ed. *Action Knoweldge and Reality*, The Bobbs-Merrill Company, Indianapolis: 295-348.

Sellars, Wilfrid, 1969, "Language as Thought and as Communication", in *Philosophy and Phenomenological Research*, 29, 4: 506-527.

Sellars, Wilfrid, 1967, *Science and Metaphysics: Variations on Kantian Themes*, Routledge and Kegan Paul, London.

Sellars, Wilfrid, 1963a, "Empiricism and the Philosophy of Mind", in *Science, Perception and Reality*, Ridgeview Publishing, Atascadero: 119-179.

Sellars, Wilfrid, 1963b, "Some Reflections on Language Games", in *Science, Perception and Reality*, Ridgeview Publishing, Atascadero: 300-333.

Sellars, Wilfrid, 1963c, "Philosophy and the Scientific Image of Man", in *Science, Perception and Reality*, Ridgeview Publishing, Atascadero: 7-40.

A cybernetic theory of persons: how Sellars naturalized Kant

Carl B. Sachs

Abstract: I argue that Sellars's naturalization of Kant should be understood in terms of how he used behavioristic psychology and cybernetics. I first explore how Sellars used Edward Tolman's cognitive-behavioristic psychology to naturalize Kant in the early essay "Language, Rules, and Behavior". I then turn to Norbert Wiener's understanding of feedback loops and circular causality. On this basis I argue that Sellars's distinction between signifying and picturing, which he introduces in "Being and Being Known," can be understood in terms of what I call cybernetic behaviorism. I interpret picturing in terms of cycles of cybernetic behavior and signifying in terms of coordination between cybernetic behavior systems, or what I call triangulated cybernetic behavior. This leads to a formal, naturalistic understanding of personhood as the capacity to engage in triangulated cybernetic behavior. I conclude by showing that Sellars's thought has the resources, which he did not exploit, for introducing the concept of second-order cybernetics. This suggests that Sellars's philosophy of mind could be developed in the direction of autopoiesis and enactivism.

Keywords: Wilfrid Sellars; Immanuel Kant; cybernetics; behaviorism; enactivism.

0. *Introduction*

It is one thing to say that we can or should "naturalize Kant", but quite another to specify in any detail what (if anything) that means – not least of which because the very phrase seems to be a contradiction in terms. Nevertheless, there is also a surprisingly long tradition of philosopher-scientists who aspired to do exactly this, beginning at least with early neo-Kantians such as Helmholtz. I do not think it controversial to suggest that Wilfrid Sellars belongs to this tradition, though it may be controversial to suggest that Sellars not only attempted to naturalize Kant, but to a remarkable extent that has not yet been fully appreciated, actually succeeded in doing so.

The linchpin of my interpretation relies on what Michael Friedman (2001) calls "philosophy as metascience". On Friedman's account, one important role for philosophical speculation is the generation of new candidate explanatory

philinq X, 1-2022, pp. 97-124
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.389
Submitted: May 2021
Accepted: October 2021

frameworks during a Kuhnian scientific crisis.[1] I suggest that Sellars's philosophy of mind should be read as a *metascience of mind* during an interregnum between behaviorism and cognitive science, when anomalies within behaviorism were accumulating and the gathering trends that would become the cognitive revolution had not yet fully congealed. Yet Sellars makes extensive use of the history of Western philosophy, from Plato through the moderns to Kant, Hegel, pragmatism, and positivism for the resources his metascience of mind requires. Hence what follows is but a preliminary sketch of how Sellars's metascience of mind established some conceptual foundations of cognitive science by translating key insights of Kantian transcendental psychology into a behavioristic-cum-computational register.

In what follows, I shall begin Sellars's first attempt at "naturalizing Kant" in his "Language, Rules, and Behavior" (1949), which turned on a remarkable and very suggestive synthesis between Edward Chace Tolman's "cognitive behaviorism" and an emphasis on "symbolic activity" that has a strongly Kantian flavor (§1). This will be followed by a somewhat longer explication of cybernetics, which has a significance for Sellars that is unfortunately almost universally neglected, and as a result of which his centrally important concept of picturing has been (I shall contend) misunderstood (§2). On this basis I will suggest a somewhat novel interpretation of Sellars's contributions to semantics and philosophy of mind (§3) before remarking on the extent to which Sellars retains any contemporary relevance to philosophy of cognitive science (§4).

One of the interesting features of Sellars's philosophy that can be brought out more clearly through a careful analysis of his engagement with the sciences of his time is his complex view of intentionality. In Haugeland's (1998) famous baseball metaphor of the positions about intentionality, he notes that there is an intermediate position between "second-base" neobehaviorism and "third-base" neopragmatism. About this, however, he says only: "Wittgenstein may have been a short-stop".[2] On the reading of Sellars I develop here, Sellars

[1] "Science, if it is to continue to progress through revolutions, therefore needs a source of new ideas, alternative programs, and expanded possibilities that is not itself scientific in the same sense – that does not, as do the sciences themselves, operate within a generally agreed upon framework of taken for granted rules. For what is needed here is precisely the creation and stimulation of new frameworks or paradigms, together with what we might call meta-frameworks or meta-paradigms – new conceptions of what a coherent rational understanding of nature would amount to – capable of motivating and sustaining the revolutionary transition to a new first-level or scientific paradigm" (Friedman 2001: 23).

[2] In the game of baseball as played in the United States, the short-stop is a defensive position located between the defensive positions at second base and at third base. For this reason, Haugeland uses the short-stop as a metaphor for a theoretical position between neobehaviorism ("second base") and neopragmatism ("third base").

develops the short-stop position into a sophisticated and defensible view because he shows how to affirm *neopragmatism* about intentionality with respect to the manifest image and *neobehaviorism* about intentionality with respect to the scientific image – a contrast that he articulates in his distinction between "signifying" and "picturing".[3]

1. *Symbolic activity as cognitive behavior*

To assess the importance of Sellars's philosophy of mind as the metascience of psychology, I want to begin where Sellars himself began: with a retrospective analysis of the debates over psychologism. These debates played a significant role in the formation of philosophy as an academic discipline, including the establishment of both phenomenology and logical positivism, both of which were formative influences on Sellars.[4] In an early text (Sellars 1947/2005a) Sellars begins by announcing that the founding move of analytic philosophy is to distinguish philosophical from psychological problems, and though he never rejected the need to distinguish normative from empirical concepts, he was also, consistently, concerned to establish the legitimacy of this very distinction.

I would like us to pay careful attention to how Sellars takes up, in the late 1940s, once again the problem or question of "psychologism". Although Sellars does not explicitly notice the connection, it is worth stressing that the critique of psychologism took for granted a specific conception of psychology itself: experimental introspectionist psychology in the grand tradition of Wundt, Titchener, and others. But we must notice (even if, perhaps, Sellars himself did not) that not all of the original arguments against psychologism can proceed once the paradigm of psychological research is no longer introspectionist but behavioristic. For example, Frege's complaint that publicly valid assertions or thoughts cannot be reduced to private mental episodes does not work when the objects of psychological research are themselves publicly observable behavioral acts. If there is something importantly right about the critique of psychologism, it must nevertheless be substantially revised in order to be applicable to behaviorism. And this is in effect what Sellars sets out to do.

In this early and unpublished note entitled "Psychologism", Sellars begins by articulating, in Kantian terms, the basic problem of the fate of epis-

[3] A comprehensive treatment of how Wittgenstein and Sellars differ in how they occupy the short-stop position is beyond the scope of this essay.

[4] For the role of psychologism in shaping philosophy as an academic discipline, see (Kusch 1995).

temology at the midpoint of the 20th century.[5] The classical conception of epistemology was by this time beleaguered on two sides by well-respected and prominent campaigns aimed at overcoming epistemology in the traditional sense: logical positivism and pragmatism. Logical positivism aimed at replacing epistemology insofar as they pursued a radical anti-psychologism that transformed epistemology into the logical analysis of science. What had been epistemology became, in the hands of the Vienna Circle and by their own admission, analytic *a priori* assertions – and hence, on the Tractarian account of analyticity that they also accepted, tautologous assertions. If the logical positivists replace epistemology with the tautologies of logical analysis, then perhaps, Sellars ventures, we should side with the pragmatists like John Dewey. Here, in sharp contrast to the anti-psychologism that shaped the context within which logical positivism emerged, we find an extremely sophisticated biologically grounded social psychology of scientific inquiry. Taking up the project developed in Dewey's *Logic* would also be a rejection of epistemology as classically conceived. The question, then, is whether there is a way of avoiding the replacement of epistemology by either logic or by science – that is, whether we could somehow salvage the very idea of synthetic *a priori* assertions, and with it, the distinct status of epistemology as not something that could be replaced by the analytic *a priori* assertions of logic or by the synthetic *a posteriori* assertions of psychology.

Though Sellars was already by 1949 tentatively sketching out the revival of the Kantian alternative to both positivism and pragmatism, he nevertheless understood the imperative of undertaking a careful examination of what positivism and pragmatism had contributed to epistemology, especially with regard to the whole question of "psychologism." To assess the viability of the arguments against psychologism, and therefore to examine whether epistemology as a synthetic *a priori* enterprise could even be vindicated, Sellars needed to first carry out a careful construction and analysis of the most sophisticated (at the time) psychological explanation of our cognitive capacities. It is only by considering what is missing from the most sophisticated psychology of cognition that we would be in the right position to vindicate the need for a genuinely *a priori* element to epistemology. At the same time, however, Sellars accepts that we need, somehow, to reconcile Kant's emphasis on the *a priori* with Dewey's naturalism: we need to sketch an *Aufhebung* between Kant and Dewey.

The next major stage of Sellars's work in which he undertakes the synthesis of Kant and Dewey is in his "Language, Rules, and Behavior" (hereafter

[5] This text is now published as an "Appendix" to Olen 2018.

LRB).[6] The initially stated goal of this text is "to explore from the standpoint of a philosophically oriented behavioristic psychology the procedures by which we evaluate actions as right or wrong, arguments as valid or invalid, and cognitive claims as well or ill grounded" (211) – in short, we are to construct a "psychology of the higher processes" that makes contact with the structure of normativity as such, whether ethical, logical, or epistemic norms. That is, we are to begin where Dewey left off, with a naturalistic, behavioristic psychology, and construct a bridge that allows us to make contact with Kantian concerns. We cannot determine by mere intuition whether or not normativity can be naturalized; we can only determine whether normativity can be naturalized by attempting to naturalize it and then evaluating whether or not the attempt is successful.

What Sellars aspires to here is a *via media* between "rationalistic apriorism" and the idea that "all meaningful concepts and problems belong to the empirical or descriptive sciences". That is, we want to avoid "descriptivism" –a tendency into which pragmatism tends to lapse – while also avoiding "rationalistic apriorism" and its accompanying "pseudo-psychology of cognitive givenness". Thus, on the one hand we must reject the error at the very heart of rationalism: the pseudo-psychology on which it depends. It depends on the error that one can, through an act of mere noticing, of the sort that would be nicely botanized by introspectionist psychologists, come to awareness of the basic underlying structures of the world (or indeed of the mind itself). In calling the cognitive given a "pseudo-psychology" Sellars is implicitly relying on how behavioristic psychologists would evaluate introspectionist psychology.[7] Yet on the other hand, "a sound pragmatism must reject descriptivism in all areas of philosophy, and that it can do so without giving one jot or tittle to what has so aptly been called the New Failure of Nerve" (213). Here Sellars is referring to Sidney Hook's article of that title in *Partisan Review* (1943), where Hook uses this phrase to refer to the tendency amongst those otherwise committed to a secular, scientific worldview to give in too readily whenever a need for pious reverence for eternal verities is announced. Thus, Sellars is explicitly aligning himself with Hook (who edited the volume in which LRB first appeared) and with Dewey (in whose honor the volume was written) while at the same time cautioning pragmatism not to reject all of the insights of the rationalism that it had come to oppose.

[6] Originally published in 1950. All page references are to the 1980 reprint in *Pure Pragmatics and Possible Worlds* edited by Jeffrey Sicha.

[7] For a behavioristic polemic against introspectionism, and one that perhaps influenced Sellars, see (Tolman 1932: 233-234).

To advance the rapprochement between rationalism and pragmatism, Sellars admits that he needs to address the philosopher's concern that psychology is not even relevant to philosophy. Why does the philosopher need to be concerned with a psychologist of symbolic behavior? "What would be the relevance of an adequate empirical psychology of rule-regulated symbol activity to the task of the philosopher?" (218). If one were to insist that the philosopher and the psychologist are engaged in different enterprises, why should the philosopher pay attention to the psychologist? The answer is that "bad psychology may give aid and comfort to bad philosophy" (*ibid.*) – that is, when we are correcting bad philosophy, we should notice how much it depends on bad psychology. But we cannot do this unless we have at least a passing acquaintance with good psychology, especially with what scientific psychology might come to say about the higher processes. In short, we will not know what we shall need to say when doing epistemology until we know more about the conceptual resources that are missing from the cognitive psychology of rational behavior. Though my use of the term cognitive psychology is anachronistic, a closer reading of the text suggests that this is precisely Sellars's concern. For his purposes, it will not suffice to carve the distinction between epistemology and psychology as between the higher, more sophisticated processes and those more primitive behaviors that we share with other animals: "To content oneself with glib phrases about stimulus-response conditioning is to give the rationalist armor and armament. … It is easy to shape the psychology of the higher processes as embodied in common sense into the direction of intuitionism and rationalism. Philosophers have been doing just that for over two thousand years" (220). For this reason, the embattled empiricist has urgent need of "an adequate psychology of rational behavior" (*ibid.*).

The most important philosophical function of "an adequate psychology of rational behavior" – of cognitive psychology *avant la lettre* – is to help the pragmatist philosopher overcome the bad "pseudo-psychology of cognitive givenness" on which rationalism and intuitionism have traditionally relied. On that model, the mind has, as it were, a *single* kind of cognitive relation: it can directly apprehend the objects referred to by terms occurring in syntactico-semantic structures (sentences, theories). Thus, one apprehends abstract entities of all sorts – universals, generals, kinds, etc. – in exactly the same way that one apprehends physical objects described by the common and proper sensibles. Sellars raises several objections to this "pseudo-psychology" over the course of his work, but I want to focus on what I shall call *the circularity objection.* The circularity objection hinges on the following thought: in order to *begin* to apprehend abstracta or universals, we would need to be able to *notice* them. But we cannot notice them without having the requisite concepts. But according to

this pseudo-psychology, the requisite concepts are directly apprehended. Thus, we cannot directly apprehend abstracta or universals, as abstracta and universals, *unless we already have*. Put otherwise, the advocate of the Given cannot avoid a "dormitive virtue" pseudo-explanation, and that is why the psychology of givenness is a pseudo-psychology.

The beginning of an alternative to the introspectionist pseudo-psychology on which rationalism depends lies in taking seriously behavioristic psychology, beginning with the thought that "most if not all animal behavior is tied to the environment in a way in which much characteristic human behavior is not" although learned habits of response "remain the basic tie between all the complex rule-regulated symbol behavior which is the human mind in action, and the environment in which the individual lives and acts" (*ibid*., 217). Crucial here is the naturalistic conviction that we are to envision the human individual as an animal in an environment, although we can consider the environment to be "social" as well as "physical". But what, exactly, does Sellars have in mind by "animal behavior" here? Although Sellars refers to behavioristic psychology in general terms, there is one specific reference that deserves closer scrutiny: the idea of a cognitive map.

Shortly before Sellars wrote LRB, the American psychologist Edward Chace Tolman published what was to become a foundational text in the transition from behavioristic to cognitive psychology: "Cognitive Maps in Rats and Men" (1948).[8] Here Tolman summarizes experiments on maze learning in rats, carried out by his graduate students and himself, to show that, contrary to the widespread view of animal behavior at the time, animal learning cannot be explained exclusively through reward-driven associations. Rather, he argued, we need to think of animals as having a map-like model of their environments that they are testing against experience and revising as necessary in order to achieve their goals and satisfy their needs. Animal behavior is not only purposive (as Tolman argued in his 1932 text) but genuinely cognitive. Hence, I shall follow Baars (1986) in referring to Tolman's position as "cognitive behaviorism", though this is not a term that Tolman himself used.[9] Though I do not mean to marginalize the importance of naturalized teleology for Tolman's pur-

[8] Tolman uses the concept of a map for the methodology of science as early as 1932, which he seems to have borrowed from his friend the pragmatist philosopher Stephen Pepper; see (Tolman 1932: 424-426). The innovation represented by the 1948 paper is that maps are not only a metaphor for scientific theories but also an analogy for animal (and human) cognition generally.

[9] There was – and remains – a lively debate as to whether Tolman was committed to realism about cognitive maps or accepted them on merely instrumentalist grounds. Though a fascinating chapter in the history of cognitive psychology, exploring it is beyond the scope of this paper. However, I believe that Sellars's own philosophy of science commits him to realism about cognitive maps regardless of the best interpretation of Tolman.

posive behaviorism (as he called his view), for present purposes I want to focus on importance of cognitive processes. On my reading, Tolman's cognitive behaviorism inspires Sellars to imagine an adequate psychology of the higher processes: one that begins with cognitive behaviorism and tries to explain *rational* behavior in terms of *cognitive* behavior.

The second prong in Sellars's strategy is to think of our symbolic activity as essentially rule-governed or rule-regulated. Here too Sellars is treading on familiar ground he has inherited from Charles Morris on signs, the Wittgenstein of the *Blue and Brown Books*, and what he learned of Cassirer from (at least) Langer's translation of *Sprach und Mythos*. What matters most to Sellars about this kind of activity is that is, in a sense difficult to articulate precisely, "free" activity – which is not to say that it is "uncaused" but rather to say that (1) it is concerned with imagining or conceptualizing non-actual possibility, and indeed with different kinds of possibility (logical, mathematical, physical), which is crucial to counterfactual reasoning and experimental testing, and also (2) the constraining rules of symbolic activity are *themselves* grounded in our acquired but revisable commitment to those rule. We can revise those normative constraints themselves – not by abandoning all rules, but by changing one rule for another. Hence our "rule-regulated symbolic activity" includes the intellectual summits of Einstein, Leibniz, and Cantor: the freely undertaken construction of new domains of syntactical and semantic structures through which our comprehension is enlarged and transformed.

The distinction between "tied behavior" – habitual responses to the environment – and "symbolic activity" – rule-regulated symbolic structures that comprise our intellectual life – is the opening move in the critique of the pseudo-psychology of cognitive givenness. As Sellars understands the state of play, the rationalist has the advantage over the naturalist for their emphasis on the inspiring intellectual achievements made in mathematics and science – but the naturalist has the advantage over the rationalist for diagnosing the cognitive given as a pseudo-psychology, the Achilles' heel of rationalism. The alternative, which Sellars emphasizes is little more than a promissory note (or at least it was in 1948), replaces the single-function account of intentionality or mindedness with a dual-function account. The crux of this account, it should be emphasized, is not simply the distinction between "tied behavior" and "rule-regulated symbol behavior" – after all, even the rationalist who has read Watson would allow for that much. Rather, what matters is that these two kinds of behavior are inextricably meshed together. If symbolic activity were not meshed together with tied behavior, it would have no causal hook-up to the environment and consequently it would be wholly irrelevant to both perception and action. If not for its meshing together with tied behavior, symbolic

activity could neither structure sensory input in the form of observations nor structure motor outputs in the form of volitions. In the absence of structuring both observations and volitions, symbolic activity would be idle if it were innate (since it could not affect perception and action) and unacquirable if it were not (since no one could learn it from observation and imitation).

What, then, does the meshing together of tied behavior and rule-regulated symbol behavior require? As Sellars sees it, "in order for the above mentioned meshing of rule-regulated language with tied symbol behavior to take place, *certain intra-organic events must function as symbols in both senses, as both free and tied symbols*" (220). That is, we need to posit neurological events – or at least neurological/non-neurological biological events – that can function as both (1) belonging to a system that coordinates purposive responsiveness to the ambient environment and (2) belonging to system characterized as a syntactico-semantic structure constituted by its own logical and material rules of inference. Let us call these *hinge events*.[10] In other words, we need to replace the single-function model of the rationalist with a dual-function model, as long as we understand that there must be hinge events: some neurological events must participate in both cognitive functions in order for them to remain coordinated (however loosely) sufficient for symbolic activity have causal bearing on the world in perception and action.[11]

Thus far I have argued for the important role of LRB in Sellars's search for an *Aufhebung* of rationalism and pragmatism, looking to both Kant and to Dewey for inspiration and guidance (among many others). The account offered in LRB is, however, a promissory note in several notable respects. In order to contextualize the route that Sellars's thought took subsequent to LRB, I want to underscore two crucial issues that Sellars neglects in LRB. First, though Sellars introduces the concept of a cognitive map and suggests that symbolic activities (including but not limited to logic, mathematics, and science) can be transposed into a naturalistic framework by seeing them as tools for constructing better cognitive maps much like those posited by Tolman, he does not articulate any causal mechanism whereby cognitive maps can be constructed and revised – without which, Sellars's naturalization of rationalism must be half-baked by his own lights. Second, Sellars does not fully articulate how we should think about the relation between the ineliminably normative

[10] The distinct status of hinge events is resumed in Sellars's much later discussion of "natural-linguistic objects" in *Naturalism and Ontology.*

[11] It is also true that the dual-purpose model is crucial to Sellars's nominalism, and it allows him to say what the rationalist wants to say about universals or kinds without a commitment to a non-naturalistic metaphysics. But while this is a strength of the Sellarsian view – if one endorses metaphysical naturalism – I shall treat it as a corollary rather than an objection.

and *a priori* nature of epistemology and the "adequate psychology of rational processes" that LRB has begun to sketch. Both of these issues occupied much of Sellars's subsequent philosophical development. I shall argue that the solution to both of these problems can be found in his mature conception of the distinction between signifying and picturing, especially in the version of that distinction that Sellars develops in "Being and Being Known".

2. *The scientific image of intentionality*

At the end of "Being and Being Known" (hereafter BBK) Sellars remarks that "recent cybernetic theory has begun to shed light on how cerebral patterns and dispositions picture the world".[12] This remarkable claim tells us that Sellars sees a deep connection between his account of picturing and what was once called cybernetics. Much like behaviorism, cybernetics has been largely forgotten because the revolution that it began has become mainstream (even though, in both cases, some of the deepest insights were forgotten along the way). What began as the science of "control and communication in animal and machine" – the subtitle of Wiener's 1948 monograph-manifesto – relatively soon evolved into computer science, information theory, and AI. Ironically, by the time that Sellars started making substantive use of cybernetic ideas, it was already beginning to be eclipsed as a serious science. For this reason (among others) the importance of cybernetics for Sellars's philosophy of mind has been, until recently, wholly neglected. Yet I shall argue that a better understanding of cybernetics is the key to Sellars's scientific image of intentionality, what he calls "picturing".

The term "cybernetics" was coined by the American mathematician-philosopher Norbert Wiener from the Greek word "kubernetes", a steersman or helmsman on a boat. The basic idea of cybernetics at the time was to refer to what were also called, at the time, "teleological mechanisms," or mechanisms capable of self-governance or self-control. An exceptionally crude precursor of such systems is the Watts governor used in steam engines. The Watts governor enables the steady production of heat by preventing too much heat from being produced: when the system overproduces, the governor closes off the supply of fuel until the pressure has decreased. The invention of electronic relays in the 20th century obliged engineers to design circuits with feedback loops so that noise can be filtered out and signals amplified relative to noise – at the same time mathematicians needed to develop a sophisticated analysis of the very concepts of "information" and "noise" that were of concern to en-

[12] Originally published in 1960; all citations to reprinted as 1963a.

gineers. Cybernetics was born from the need to conceptualize, operationalize, and realize the concepts central to information theory, computer science, and their sequelae.[13]

The crucial notion that Sellars absorbs from cybernetics is the idea of feedback. In Wiener's formulation, feedback is indispensable "when we desire a motion to follow a given pattern the difference between this pattern and the actually performed motion is used as a new input to cause the part regulated to move in such a way as to bring its motion closer to that given by the pattern" (Wiener 1948: 6-7).[14] In light of this, feedback is essential to patterned behavior in general: patterned behavior is possible due to feedback that corrects actual deviations, errors, or noise relative to what expected or desired. In the case of designed systems, it is the designers who know what pattern they want to see generated and institute feedback loops in order to generate the behavior that they intend. In the case of naturally evolved cognitive systems, there is no designer, and yet can say that patterned behavior emerges from the feedback loops between the cognitive map (as a spatio-temporal map of the environment and the place of the organism in that environment) and the environment to which that map is structurally coupled via transducers and effectors. The concept of feedback is also crucial here because it allows us to understand how Sellars transforms the concept of picturing that he has borrowed from Wittgenstein's *Tractatus.* Put much too simply, Sellars uses the concept of feedback to give Tractarian picturing a cybernetic twist.[15]

A corresponding change in the basic metaphysics is required by the new science of cybernetics, since we now must understand change not only in terms of energy but also in terms of the then-new concept of information:

> the newer study of automata, whether in the metal or in the flesh, is a branch of communication engineering, and its cardinal notions are those of message, amount of disturbance or "noise" – a term taken over from the telephone engineer – quantity of information, coding technique, and the like. In such a theory, we deal with automata effectively coupled to the external world, not merely by their energy flow, their me-

[13] See Kline 2017 for the history of cybernetics, but especially the personal and political factors that led to its eclipse. In large part cybernetics was re-branded as information theory and as computer science; it also led directly to chaos theory, complexity theory, evolutionary robotics, autopoiesis, artificial intelligence, and cognitive science.

[14] I am focusing on Wiener partly because of his historical importance and partly because Sellars had a copy of Wiener's *Cybernetics* (1948) in his personal library, though it was not his only source of information about cybernetics.

[15] This becomes the key move in how to understand rule-regulated behavior – a rule is a generalization that tends to make itself true by virtue of how the norm is enacted through feedback loops between members of the community.

tabolism, but also by a flow of impressions, of incoming messages, and of the actions of outgoing messages. (Wiener 1948: 42)

Of particular interest to the history of the scientific image of mind is how Wiener insisted on conceptualizing the central nervous system in cybernetic terms. Two passages are noteworthy for the parallel between Wiener and Sellars:

> The central nervous system no longer appears as a self-contained organ, receiving inputs from the senses and discharging into the muscles. On the contrary, some of its most characteristic activities are explicable only as circular processes, emerging from the nervous system into the muscles, and re-entering the nervous system through the sense organs, whether they be proprioceptors or organs of the special senses. (Wiener 1948: 8)

and

> [F]or effective action on the outer world it is not only essential that we possess good effectors, but that the performance of these effectors be properly monitored back to the central nervous system, and that the readings of these monitors be properly combined with the other information coming in from the sense organs to produce a properly proportioned output to the effectors. (Wiener 1948: 96)

As we shall see, this is precisely how Sellars characterizes the "anthropoid robot of the future" in "Being and Being Known" as having internal computational states that covary with the states of the environment and its body due to feedback loops between processors, effectors, and transducers.

It is also noteworthy, I think, to stress that Wiener regards cybernetics as bearing directly on the question as to whether logic is reducible to psychology – that is, to "psychologism":

> The science of today is operational; that is, it considers every statement as essentially concerned with possible experiments or observable processes. According to this, the study of logic must reduce to the study of the logical machine, whether nervous or mechanical, with all its non-removable limitations and imperfections. … any logic which means anything to us can contain nothing which the human mind – and hence the human nervous system – is unable to encompass. (Wiener 1948:125)

By appealing to a version of operationalism, Wiener is able to suggest that the meaning of logical statements is equivalent to the procedures used by a computing machine – whether metal or meat – used to verify those statements. This does not reduce logic to psychology, nor psychology to logic – but it does transform logic into the science of the formal properties of cognitive machin-

ery. All of this is nicely taken on board by Sellars throughout the 1950s, so that by the early 1960s, Sellars is finally in a position to use cybernetics for conceptualizing a scientific image of mind that allows him to articulate a comprehensive, philosophically adequate alternative to "the pseudo-psychology of cognitive givenness" upon which rationalism depended.[16]

In the development of this alternative, "Being and Being Known" (Sellars 1963a) deserves special status because it is here that Sellars explicitly invokes cybernetics in his scientific image of intentionality. This conception is developed through a close criticism of the Aristotelian philosophy of mind located in (among other places) Thomism. Sellars does this for two main reasons. The first is that he thinks that there are important insights in the Aristotelian tradition that have been overlooked by the modern approach that begins with Descartes. The second, and more important, is that Sellars's philosophical method suggests the following commitment: the rational defensibility of his conception depends on its place within the dialectic of the history of philosophy of mind.

Sellars suggests that we accept the Aristotelian distinction between cognitive functions as systematically related to the environment and cognitive functions as governed by rules. But how is this distinction to be understood? As Sellars sees it, this distinction is not one between kinds of cognitive function but rather between different ways of thinking about what cognitive functions are. For while we have a few thousand years of theorizing about cognitive functions using the conceptual resources of the manifest image, we can also begin to compare those theories with the account of cognitive functions using the conceptual resources of the scientific image. To do this, Sellars engages in the thought-experiment of imagining an "anthropoid robot of the future" (Sellars 1963a: 51) – something that, perhaps falling short of genuine artificial general intelligence, might be within the next generations of Mars rovers.

Sellars's starting point is to accept the traditional idea – going back at least to Aristotle – that there is an isomorphism between the intellect and the world – that *veritas est adaequatio intellectus et rei*. But he suggests that this isomorphism must be understood in two very different senses, and that nothing but confusion results from conflating these two distinct senses. The two senses refer to different "orders": "the logical order" and "the real order." The former is the explication of the order of understanding (*ratio cognoscendi*); the latter is the explication of the order of being (*ratio essendi*). In the real order, the

[16] Though Sellars lobbies this accusation in rationalism in LRB, by the time he writes "Empiricism and the Philosophy of Mind" he has realized that the myth of the given is a problem not only for rationalism and for empiricism but even for Kant and Hegel. For a brief reconstruction of the history of epistemology aimed at making sense of this claim (see Sachs 2020).

isomorphism of intellect and world is what he calls "picturing"; in the logical order, the isomorphism of the intellect and the world is "signifying". Confusion between these two orders has led to the misbegotten Platonic-Aristotelian idea that the intellect signifies the world by being informed by immaterial natures. Hence Sellarsian nominalistic materialism requires a sharp demarcation between signifying and picturing.

What we need at this point is an answer to the question, "what are we talking about when we talk about how the intellect pictures the world?" And to this the Sellarsian answer is: *cybernetics*. It is cybernetics that Sellars is alluding to when he writes, "I shall present the distinctions I have in mind as they appear when projected into discourse about computing machines, guided missiles, and robots" (Sellars 1963: 51). These are more or less standard examples in the cybernetics literature of the 1940s through 1970s:[17]

> Suppose such an anthropoid robot to be "wired" in such a way that it emits high frequency radiation which is reflected back in ways which project the structure of its environment (and its "body"). […] Suppose such a robot to wander around the world, scanning its environment, recording its "observations", enriching its tape with deductive and inductive "inferences" from its "observations" and guiding its "conduct" by "practical syllogisms" which apply its wired-in "resolutions" to the circumstances in which it "finds itself". It achieves an ever more adequate adjustment to its environment, and if we permitted ourselves to talk about it in human terms (as we have been) we would say that it *finds out* more and more about the world, that it *knows* more and more *facts* about what took place and where it took place, some of which it *observed*, while it *inferred* others from what it did *observe* by the use of *inductive generalization* and *deductive reasoning*. (Sellars 1963a: 52-53; emphasis original)

We can, from the standpoint of the electronic engineer or cybernetician, consider the states of the robot as building up a picture of the environment – although "this picturing cannot be abstracted from the mechanical and electronic processes in which the tape is caught up" (Sellars 1963a: 53), or as we might say today: cognition is *both* computational *and* necessarily embodied and embedded. Just as the grooves on a record player cannot be understood apart from the procedures by which records are produced and played, so too the computational states of the robot cannot be understood apart from the physical *habitus* of the robot.[18]

Wittgenstein also uses the record as an example of picturing at *TLP* 4.014:

[17] See Rosenblueth, Wiener, and Bigelow 1943; Wiener 1948.

[18] See Huebner 2018 for a detailed explanation for why the physical *habitus* of the robot is necessary for understanding the analog computations and analog representations that comprise the robot's "mind".

"A gramophone record, the musical idea, the written notes, and the sound-waves all stand to one another in the same internal relation of depicting that holds between language and the world" (Wittgenstein 1974: 20). The crucial difference is that Sellars uses cybernetics to correct Wittgenstein: picturing is instituted through feedback loops between the tape, the robot's body, and its ambient environment.

Sellars's invocation of computational states may seem to clearly anticipate what has become known as "the computational theory of mind", especially in the versions promoted in Putnam in mid-1960s and by Fodor in the late 1970s. However, there is a crucial difference between CTM and Sellars's cybernetics. As Sellars sees it, the computational states that comprise the mind cannot be disentangled from the whole network of behaviors in which they are embedded, as made vivid by his comparison of the mind with a vinyl record. If one were to carefully examine the surface of a vinyl record, one can discern hundreds of thousands of grooves etched into it. But in order to understand why that record has the grooves that it does, one needs to understand the record in context, both as the result of a manufacturing process whereby sounds are converted into a semi-stable form and as something that can be inserted into an audio system designed to reproduce the sounds that were translated into the record when it was manufactured. The structure of the grooves is a consequence of the transposition of the structure of the music from an acoustic medium to a vinyl medium.[19] In the same way, the computational states of the mind are a "materialization" of the features of the environment that caused those states via perceptual episodes. The key difference is that the structures are distorted or modulated at the same time that they are transposed from environment to mind, so that the relationship between them is not a simple matching but rather a highly dynamic structural coupling between computations and environmental features.[20]

[19] In one crucial respect the vinyl record analogy is misleading. The recording and production process allows for near isomorphism between the acoustic properties of the music and the grooves in the record, which is why vinyl is preferred even today by purist audiophiles. MP3s and other compression formats are comparatively quite "lossy" – there is loss of information as the signal is compressed – meaning that the mapping relation between playback and original is homomorphic, not isomorphic. Lossy compression formats are almost certainly a better metaphor for animal sensory systems than the near isomorphism of vinyl recordings (see Akins 1996). However, vinyl records are a useful metaphor because they record signals in an analog format, rather than a digital one, although we probably do not have a clear understanding of analog vs digital processing in neuronal assemblies.

[20] In other words, Sellars accepts with the cognitivists that the mind is comprised of computational states that function as representations of the environment, but he also insists, along with the anti-representationalist and 4E proponents, that cognition is necessarily embodied and embedded. (see Huebner 2018 and Sachs 2018).

I shall call this position "cybernetic behaviorism". It differs from Tolman's "cognitive behaviorism", which was already important for LRB, by explicitly drawing upon cybernetics for concepts (e.g., feedback loops) and examples (e.g., guided missiles) in theorizing about how cognitive maps are constructed and updated. Thus, while Tolman argues that intelligent, purposive behavior is best explained by positing a map-like mental model of the features of the environment constructed as the animal sensed and interacted with that environment, he did not propose any underlying mechanism. Sellars, by contrast, uses cybernetics to propose an underlying computational basis to thought driven by sensorimotor feedback loops.[21]

Unlike a more "Cartesian" form of cognitivism, Sellars underscores the importance of the physical *habitus* of the cognitive system matters – the kinds of maps it will construct is inseparable from its iterated feedback loops with ambient environments.[22] What picturing is: feedback loop driven updating of nonconceptual representational states functionally embedded in a computational information processing system that, as a dissipative structure, continually exchanges causal flows of energy-matter with its ambient environment.

The thought experiment of the BBK robot thus puts a cybernetic spin on purposive behaviorism: the robot's purposive behavior can be explained from the perspective of the electrical engineer in terms of feedback loops between two systems – the ambient environment and the robot – informationally coupled through transducers and effectors. The upshot of the thought experiment is that the cognitive friction with the environment that both rationalists and empiricists sought to explain with something Given – whether "the *illuminatio* of Augustine" or "the *data* of the positivists" (Sellars 1963b: 356) – can be explained entirely by adopting the scientific image of mind: cybernetic behaviorism.

Cybernetic behaviorism is crucial for understanding Sellars's argument for why semantic terms such as "means", "refers to," and "is about" do not designate a relation between mind and world.[23] These terms belong to the manifest image of intentionality: they are the product of millennia of philosophical reflection on the world of everyday life and experience, and they have valu-

[21] However, it was not until the 1980s that Sellars finally applied the insights from cybernetics *directly* to biological systems, and not just as an analogy with them. It was at this time that he developed what he came to call "animal representational systems." In this regard Sellars was influenced by the cognitive revolution.

[22] The exact relevance of this account to contemporary debates between cognitivism and 4E cognition depends in part on whether the relation between the computational states of the robot and its body and environment is one of coupling or constitution (see Rowlands 2010). Further exploration of this point is beyond the scope of this paper.

[23] McDowell (1998) ascribes to Sellars the position that intentionality is not a mind-world relation. For an incisive criticism of this interpretation, see (Shapiro 2011).

able roles to play in the elucidation of discourse. But although semantic terms have a "surface grammar" of designating a mind-world relation, taking them at face value inevitably leads to positing intensional entities: meanings, propositions, *Sinne*, ιδέα. Sellars's strategy for securing a nominalistic, materialistic metaphysics does not (*pace* McDowell) require him to deny that intentionality is a mind-world relation; rather, his strategy is to argue that the manifest image conception of intentionality is now replaceable by a scientific image of intentionality.[24] In the scientific image of intentionality, we retain the manifest image commitment to the idea that intentionality is a mind-world relation. The crucial difference is this: the exact nature of the mind-world relation is not based on a conceptual explication of semantic vocabulary but rather on a causal explanation of cybernetic mechanisms.

3. *Cybernetics, community, and personhood*

Based on this admittedly quick and crude sketch of what I have been called cybernetic behaviorism, I shall now develop what I take to be a Sellarsian solution to long-standing problems in the philosophy of mind – chiefly, the nature of *content*, *intentionality*, or *meaning*, which is (Sellars thinks) a problem at the very heart of what it means to be a thinking thing. Though Sellars inherits much from German Idealism and American Pragmatism with regard to the indispensable role of membership of a community in our self-conception as rational thinkers and agents, he also reworks this inheritance using cybernetic behaviorism.

The germinal seed of a Sellarsian account can be found in what he says about the conditions under which it would make sense to talk about the meaning of a machine state of the BBK robot. The states of the BBK robot, which picture its environment, are said to signify – to have meaning, to have content – insofar as we can *coordinate* our signifying behavior with its picturing behavior. Just as one can utter the English sentence "'*grun*' means *green*" to convey to an English speaker what the German speaker means by the German word "*grun*", we can also construct translation manuals for the BBK robot. We ascribe semantic content to the BBK robot to the extent that a translation manual can be constructed.[25]

[24] However, there is another sense of intentionality, "the language of individual and community intentions", which persists in the scientific image. What is replaced by picturing is the sense of intentionality that involves the "aboutness" or world-directedness of thought. I would like to thank Willem deVries for pressing me to be clearer on this point.

[25] Though I believe the Quinean term "translation manual" is not inappropriate in discussing Sellars, there are two crucial and relevant differences: the Sellarsian translation manual is not con-

To construct a translation manual, we need to be able to successfully notice what in our shared environment that the robot is responding to, classify the picturing states of the robot that are individuated by virtue of their causal role in the robot's sensorimotor feedback loops, and compare those states with our concepts as we use them in our social practices.[26] This triadic process – between us, the robot, and the environment – can be usefully conceptualized, following Davidson (1990; see also Davidson 1992), as a process of "triangulation". At the heart of Sellars's theory of meaning or conceptual content is what I will therefore call *triangulated cybernetic behaviorism*. The function of the ascription of semantic content is to facilitate triangulated cybernetic behavior: to construct a coordination device whereby we can say of our own linguistic behavior that it is similar enough to other linguistic behavior that the conditions for successful cooperation have been established.

But what, *in rerum natura*, is linguistic behavior? The Sellarsian answer is that linguistic behavior just *is* triangulated cybernetic behavior: when we ascribe semantic content to any utterance or inscription – even our own – is that it can be coordinated with other utterances or inscriptions that are functionally integrated into the sensorimotor feedback loops of other cybernetic systems, where the criteria of coordination lie in successful cooperation. How a cybernetic system interacts with its environment depends on how it models that environment, which means that cybernetic systems can cooperate only to the extent their models are sufficiently consistent that the actions guided by those models do not generate conflict. It is important to keep distinct the role of semantic attribution and the role of predictions and explanations of behavior. If you attribute to me the belief that geese are ducks, you are both making a claim about how I picture waterfowl and hence how I will engage with them and also making a claim that my picturing is incompatible with picturing based on sound scientific taxonomy.[27]

What we need at this point is an account of how we are to understand the relation between the manifest image of intentionality as embedded in our folk psychology and the scientific image of intentionality as explicated in cybernetic behaviorism, including triangulated cybernetic behavior. Though Sellars returns to this problem throughout his work, I want to focus on how he thinks about it at the same time as he is developing his theory of meaning in conjunction with cognitive/cybernetic behaviorism.

structed through stimulus-response pairs and it does not neglect the role of internal information processing.

[26] This is not to insist that all of the robot's states picture; if it has constructed cognitive maps that include abstract or theoretical terms, then those states are not themselves picturing, though they are required for picturing.

[27] Thanks to Willem deVries for the "Carl believes geese are ducks" example.

Sellars's claim that normative statements are logically irreducible to natural statements *and yet causally reducible* has provoked a good deal of commentary. As I see it, the crux of the argument depends on how Sellars understands meaning as functional classification. To say that mind is "logically irreducible" to body is to say *only* that the class of analytic truths does not include statements that stipulate an identity relation between statements made in folk psychological discourse and statements made in a suitably scientific discourse – cognitive behaviorism augmented by cybernetics. There is no equivalence of intension between statements made from within the intentional stance and statements made from within the cybernetic stance; no statement relating those statements could be true "by meaning alone".[28]

But if folk psychological statements and cybernetic behavioristic statements are *not* intensionally equivalent or synonymous, that nevertheless leaves open the possibility of co-extension. And this is not a possibility that Sellars rejects, though his acceptance of it takes a curious form: he says that folk psychological statements and cybernetic behavioristic statements "convey the same information." This should give us pause, because in 1953 the very idea of "information" as something that could be "conveyed" was just beginning to coalesce; Shannon's probabilistic definition of "information" was only published in 1948, five years earlier. John O. Wisdom's "The Hypothesis of Cybernetics" appeared in 1951, and that was one of the first philosophers to take up cybernetics.[29]

The relation between intentionality and cybernetic behavior is, however, slightly more complex than this suggests. From one perspective – that of the scientific image under construction – cybernetic behavior is the scientific image of thought. But if that were the end of the matter, what would become of the classical conception of intentionality as semantic content that is about the world in which it used? Sellars's answer to this question depends not just on his analysis of normativity but also on the role of that analysis in his understanding of community.

Sellars accepts and develops the classical German Idealist emphasis on the ineliminable normativity of rational thought and action: intentionality and normativity are logically interdependent. Cybernetic behaviorism cannot suffice for the scientific image of intentionality unless it can somehow accommo-

[28] Sellars never accepted Quine's critique of the analytic/synthetic distinction or the implications of that critique for ontological commitment, because Sellars – unlike Quine and, for that matter, Carnap – never thought that we should reject intensional semantics, though he was sympathetic to Morton White's view that the analytic and the synthetic had become an untenable dualism. Sellars's response to White is to repair the distinction, not reject it.

[29] John O. Wisdom, a philosopher of psychology, is not the same person as the ordinary language philosopher John Wisdom, though they were cousins.

date this classical emphasis on normativity – and indeed, not just on normativity *simpliciter* but on the close tie between normativity and sociality. The rules of criticism that govern language-entry transitions (perceptions), formal and material inferences, and language-exit transitions (volitions) are, for us rational animals, interlocked with rules of conduct whereby we hold each other accountable for what we claim to perceive, think, and do.

To the extent that cybernetics could perhaps explain the rules of criticism or ought-to-be rules that govern the lives of non-rational animals, it would be in a weak or analogical sense – since those animals are (*ex hypothesi*) incapable of regarding themselves as governed by rules of criticism, it is we who regard them as being so governed, with the rules themselves being a consequence of how past natural selection has shaped the ways in which those animals occupy their niches (however plastic). In these cases, what we describe as rules of criticism are explained in terms of feedback across brain-body-environment causal loops, where the issuing of the rules functions as negative feedback to prevent behaviors that deviate too much from the rules.

Even if something like this were made plausible – and there is little doubt that Sellars himself thought so – Sellars would certainly accept that the story for us rational animals cannot be quite that simple, and that is because what distinguishes us *qua* rational animals is not just that we can regard ourselves as governed by rules of criticism but also that we regard ourselves as being so governed by virtue of the interlocking relationship between rules of criticism and rules of conduct. When a little brown bat fails to capture a fleeing mosquito, it has done something that ought not be the case about what bats do – it has, in a broad sense, made a mistake – but it has not transgressed against Chiropteran social practices, for there are none. By contrast, when a person looks at an alligator and calls it a crocodile, they have not used the words correctly and are susceptible to correction from others.

This line of thought suggests that the scientific image of mind based on cybernetic behaviorism will be incomplete unless it can somehow accommodate not only rules of criticism but also rules of conduct. Without an account of normativity, the scientific image would be radically incomplete – it would not be a scientific image *of mind*. Thus, what we require here is an account that yokes together what Sellars says about the ineliminably normative dimension of human thought and action, based as it is on the philosophical clarification and elucidation of the manifest image, with what cybernetic behaviorism says about the scientific image of cognition.

In these terms, what are we to say about the ineliminable role of rules or norms in our linguistic and non-linguistic social practices? If the ascription of semantic content is to convey that the success (or failure) of triangulated cyber-

netic behavior, then the utterance or gestures of norms or rules that underpin meaning ascription are the behaviors that bring about that coordination. Rules or norms are ineliminable because rationality – or at least the human form of rationality – is necessarily social.[30] Social life does not require perfect or ideal cooperation – at least not to a degree that would eliminate all conflict – but it requires that cooperation be, if not optimal, at least satisficing enough of the time for social life to be reproduced from one generation to the next.

With this element in place, we can finally draw out the following implication for a Sellarsian theory of personhood. In the concluding paragraphs of PSIM, Sellars remarks that to say of something – whether "a featherless biped or a dolphin or a Martian" (Sellars 1963c: 39) – that it is a person is to say that it is a member of one's community. It is to say that the naturalistic basis of community is triangulated cybernetic behaviorism: personhood is the status of a cybernetic system that actualizes a capacity to triangulate its behavior with other cybernetic systems that can also actualize their capacities for triangulated behavior. Triangulated cybernetic behavior is realized via the interlocking relation between rules of conduct and rules of criticism, such that it can say (or think) "I am one of you".

4. *Sellars, cognitivism, and enactivism*

A careful examination of the importance of cybernetics for Sellars's philosophy of mind has substantial implications for how we should assess his thought in light of contemporary cognitive science. This is because cybernetics is the ancestor of both cognitivism, with its emphasis on cognition as rule-governed manipulation of symbolic representations, and enactivism, with its emphasis on non-representational dynamic coupling between biologically autonomous systems and their environments. For this reason, I want to briefly explore how cybernetics came to influence both cognitivism and enactivism before indicating a place in Sellars's thinking where he could have re-oriented his ideas in a more enactivist direction than he actually did.

The rift between cognitivism and enactivism can be traced, according to Froese (2010), to the emergence of the split between computer science and second-order cybernetics. The decisive issue turned on how cyberneticians responded to Ashby's demonstration that seemingly intelligent complex behavior could emerge from purely mechanistic assemblages. This called into question the presumptive realism at the heart of cybernetics as an objective science:

[30] However, a correct Sellarsian reading of this point may require separating "the human" as normative concept from *Homo sapiens* as a biological concept (see Wolfendale 2019).

if we ourselves are just machines turning "noise" into "meaning", then what are the rational credentials of any "output" from the electrochemical computer called the brain?

This Ashbyian crisis, as Froese calls it, elicited two responses from the cybernetic community. The first response, which became computer science and cognitivism, remained committed to a realist epistemology and sought to implement it mechanistically by treating cognition as the mechanistic manipulation of symbolic representations. Consequently, there was no need to ground symbolic representations on anything more basic or mechanistic: the manipulation of symbols *was* cognition. The second response, which became second-order cybernetics and enactivism (among other paradigms), rejected the realism that defined cybernetics and instead embraced a constructivist epistemology. As Heinz von Foerester came to put it, the job of the brain is to compute an effective model of reality.That the enactive and autopoietic approaches to cognition developed out of second-order cybernetics is relatively well-known (Froese 2010, Froese 2011). The basic idea of second-order cybernetics, in von Foerster's terms, turns on a shift from "observed systems" to "observing systems". In first-order cybernetics we are describing circular causality – recursion or feedback loops – in systems that we have built: we study them as objective components of the material universe. In second-order cybernetics we are describing circular causality in the systems that *we ourselves are.* But because we are persons, members of communities structured by relations (and asymmetries) of recognizing and being recognized, second-order cybernetics had to relinquish the commitment to pure objectivity: in becoming part of the conceptual structure in which we experience and understand ourselves and others, it became necessary for the first-personal and second-personal perspectives to enrich the cybernetic vocabulary.

I want to now suggest a Sellarsian argument in support of second-order cybernetics. Briefly put, Sellars's criticism of the Given is best understood as rejecting the idea that how we experience the world can be decisively and clearly demarcated and protected against changes in our conceptual structure as a result of new discoveries in the empirical and formal sciences.[31] Despite his appreciation for phenomenology, Sellars rejects the Husserlian idea that the life-world can or should be defended against incursions by the sciences. Rather, for Sellars, the goal of joining the scientific and manifest images requires *incorporating* the sciences *into* the life-world – and this is precisely what second-order cybernetics does.

[31] See O'Shea 2021 as to why a commitment to the Given is a commitment to holding that cognitive experience has a categorial structure that is unrevisable, come what may.

In other words, by incorporating cybernetic concepts into how he understood himself and others, von Foerster executed a Sellarsian strategy for incorporating scientific concepts into the world of everyday life. As Sellars puts it, "by construing the actions we intend to do and the circumstances in which we intend to do them in scientific terms, we *directly* relate the world as conceived by scientific theory to our purposes, and make it *our* world and no longer an alien appendage" (Sellars 1963c: 40). But we could not do this if the Given were not a Myth. For if the Given were not a Myth, there would be a clearly discernible stratum of our experience that would be unrevisable, come what may any changes elsewhere in our conceptual structure; it would not be possible to observe oneself or others as cybernetic systems.

This point can also be framed in terms of "the Myth of Jones" at the concluding sections of "Empiricism and the Philosophy of Mind" (Sellars 1963d: 183-196). There, Sellars imagines a group of people – "our Rylean ancestors" – who lack the concepts of *thought* and of *sensation*. The concepts are invented by a mythical "Jones" who one day, puzzling over certain behaviors – that people act as if they had been talking to themselves but without saying anything aloud, or that people act as if they are seeing or hearing things that no one else sees or hears – comes up with the concepts of thought and of sensation by analogy. People have thoughts that are like overt verbal episodes, except that no one can hear or see them; and they are sensations that are like the sensible qualities of physical things, except that they only exist for that person, in their "consciousness". Over time, Jones teaches these innovations to others, so that what began as a theoretical posit becomes part of our non-inferential awareness of self and others. Transposing this lesson from thoughts and sensations to cybernetics, we can say that Heinz von Foerster was the "Genius Jones" of cybernetics.[32]

Thus far I have only suggested that the transition from first order to second order cybernetics is one that makes sense philosophically in light of the criticism of the Myth of the Given. But I think that we also need to take careful notice of two further considerations: how the concepts of second-order cybernetics become transformed as a consequence of being incorporated into the lifeworld, of first-person and second-person linguistic performances, and how this transformation affects our reading of Sellars himself.

On the first point: first-order cybernetics depended essentially on abstracting away from the differences between organisms and machines, in order to produce abstract concepts like "information" and "feedback". Neglecting the material reality of physics, chemistry, and biology was necessary for producing

[32] I put it this way largely for rhetorical effect – I do not intend to slight the contributions of Maturana, Pask, Beer, Varela, Bateson, Mead, Thompson, and many others.

the abstract models that the cyberneticians analyzed and debated. For example, when McCulloch and Pitts (1943) demonstrated that recurrent networks of neuron-like elements can realize Boolean functions, they explicitly introduced the simplifying assumption that each neuron either does or not fire. They constructed an abstract model, a digital neuron. The fact that their model did not take into account biological reality, where neurons are modulating their activity in ways that are much more "analog" than "digital", was not among their concerns. Likewise, Wiener deliberately chose as the subtitle of his manifesto "control and communication in the animal and in the machine" – with the assumption that there is no difference that makes a difference between animals (including us) and machines (including computers).

This simplifying abstraction, however important for allowing the conceptual and empirical breakthroughs of first order cybernetics, could not be maintained when cybernetics became incorporated into how we understand ourselves: the abstract had to become concrete. It is for this reason that second order cybernetics rather quickly evolved into autopoiesis theory: a formal model about the specific kinds of organizational features that a complex system must have in order to be described as "alive". Autopoiesis theory and related approaches, such as those of Robert Rosen, Stuart Kauffman, Alvaro Moreno, Matteo Mossio, and Ezequiel di Paolo, have generated a rich and sophisticated way of understanding why organisms are *not* machines and cognition is *not* computation – *contra* both first-order cybernetics and the cognitivist research program that it also gave rise to.

On the second point: this reconstruction of the transition from first order cybernetics to second order cybernetics also matters for Sellars, because as I have argued here, his own philosophy of mind is deeply indebted to first order cybernetics. Sellars is (perhaps) the first computational functionalist in philosophy of mind, because of how he incorporates cybernetics into his understanding of cognitive systems. But, somewhat ironically, Sellars himself does not take von Foerster's step of incorporating cybernetics directly into the life-world: a step that Sellars should have taken given the larger shape of his thought, and perhaps one that he would have taken if this inconsistency had been pointed out to him.

The fact that Sellars did not take this step has had further repercussions for post-Sellarsian philosophy – that is, philosophy that takes itself to be building upon Sellars's considerable achievements. Because Sellars himself did not take the von Foerster step of incorporating cybernetics into his experience of himself and the world, he did not question the realist epistemology that first order cybernetics, like all modern objective thought, took for granted. For Sellars, the goal of science is to construct testable models of the fully determinate regularities that exist in a fully mind-independent sense, and which the phi-

losopher can use to tell us which aspects of the phenomenal world are truly mind-independent and which are not. These commitments also shape how post-Sellarsian philosophers, such as Daniel Dennett, Paul Churchland, and Robert Brandom, took up the legacy (in very different ways) of scientific realism and computational functionalism. These philosophers belong to the legacy of first-order cybernetics and cognitivism because they continue Sellars's refusal to take the step towards second-order cybernetics, even though Sellars himself could have done so – and arguably should have. Sellars's cybernetic behaviorism is grounded in his appropriation of first-order cybernetics, but in light of the rift between cognitivism and enactivism, it is questionable whether first-order cybernetics was a dialectically stable position. As I see it, Sellars's overarching project should have led him to adopt second-order cybernetics – though without necessary abandoning his commitment to representationalism, despite the anti-representationalism that has become definitive of enactivism.[33]

5. *Conclusion*

It has become something of a commonplace that Sellars's complicated image of humanity in the universe belongs to the tradition of philosophers who aspired to naturalize Kant. The extent to which Sellars succeeded in doing so has been obscured by confusion about what he meant by "picturing" – or even if we need it at all. I have tried to show that Sellars's concept of picturing had been difficult to understand due to ignorance of his historical context. With the proper context in place, we can see that Sellars drew upon Tolman's cognitive behaviorism and Wiener's cybernetics to transform Wittgensteinian picturing into a cybernetic-behavioral concept. In doing so, Sellars conceptualized picturing in terms of feedback loops as being described and built by the first generation of cyberneticists. It is because Sellars's use of cybernetics was not even fully appreciated in Sellars's own time, and has been completely forgotten since, Sellars's distinct version of naturalizing Kant has not yet received the full treatment that it merits.[34]

Carl B. Sachs
Marymount University
csachs@marymount.edu

[33] A different route from Sellars to contemporary autonomy or enactive theories could be traced by considering Sellars's own project of "naturalizing Kant" in light of the Third Critique account of teleological judgment (see Weber and Varela 2007).

[34] I would like to thank Willem de Vries, Bryce Huebner, and Kyril Popatov for their detailed and encouraging criticisms of previous versions of this paper, and comments from Evan Thompson on §4. I would also like to thank Steve Levine for our many conversations about Sellars over the years.

References

Akins, Kathleen, 1996, "Of Sensory Systems and the 'Aboutness' of Mental States", in *Journal of Philosophy*, 93,7: 337-372.

Baars, Bernard, 1986, *The Cognitive Revolution in Psychology*, The Guilford Press, New York.

Davidson, Donald, 1990, "Epistemology Externalized", in Davidson 2001: 193-204.

Davidson, Donald, 1992, "The Second Person", in Davidson 2001: 107-122.

Davidson, Donald, 2001, *Subjective, Intersubjective, Objective*, Oxford University Press, Oxford.

Friedman, Michael, 2001, *Dynamics of Reason,* Center for the Study of Language and Information, Stanford.

Froese, Tom, 2010, "From cybernetics to second-order cybernetics: A comparative analysis of their central ideas", in *Constructivist Foundations*, 5,2: 75-85.

Froese, Tom, 2011, "From Second Order Cybernetics to Enactive Cognitive Science: Varela's Turn from Epistemology to Phenomenology", in *Systems Research and Behavioral Science*, 28, 6: 631-645.

Haugeland, John, 1998, "The Intentionality All-Stars", in *Having Thought: Essays in the Metaphysics of Mind*, Harvard University Press, Cambridge MA: 127-170.

Hook, Sidney, 1943, "The New Failure of Nerve", in *Partisan Review*, 10,1: 2-23.

Huebner, Bryce, 2018, "Picturing, Attending, and Signifying", in *Belgrade Philosophical Annual*, 31: 7-40.

Kline, Ronald, 2017, *The Cybernetics Moment: Why We Call Our Age the Information Age*, Johns Hopkins University Press, Baltimore.

Kusch, Martin, 1995, *Psychologism; A Case Study in the Sociology of Philosophical Knowledge*, Routledge, New York.

McCulloch, Warren and Walter Pitts, 1943, "A Logical Calculus of the Ideas Immanent in Nervous Activity", in *The Bulletin of Mathematical Biophysics*, 5: 115-133.

McDowell, John, 1998, "Intentionality is a Relation", in *Journal of Philosophy*, 95, 9: 471-491.

Olen, Peter, 2016, *Wilfrid Sellars and the Foundations of Normativity*, New York, Palgrave.

O'Shea, James, 2021, "What *is* the Myth of Given?", in *Synthese*, 199: 10543-10567. <https://doi.org/10.1007/s11229-021-03258-6>.

Rosenblueth, Arturo, Norbert Wiener, and Julian Bigelow, "Behaviour, Purpose and Teleology", in *Philosophy of Science*, 10: 18-24.

Rowlands, Mark, 2010, *The New Science of the Mind*, The MIT Press, Cambridge MA.

Sachs, Carl, 2019, "In Defense of Picturing: Sellars's Philosophy of Mind and Cognitive Neuroscience", in *Phenomenology and the Cognitive Sciences*, 18,4: 669-689.

Sachs, Carl, 2020, "A Conceptual Genealogy of the Pittsburgh School: Between Kant and Hegel", in K. Becker and I. Thompson, eds. *The Cambridge History of Philoso-*

phy, 1945-2010, Cambridge University Press, Cambridge.

Sellars, Wilfrid, 1980, "Language, Rules, and Behavior", in Sicha, Jeffrey, ed. *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*, Ridgeview Publishing Company. Reprinted from 1950, "Language, Rules, and Behavior", in Hook, Sidney, ed. *John Dewey: Philosopher of Science and of Freedom*, Barnes and Noble, New York.

Sellars, Wilfrid, 1963a, "Being and Being Known", in *Science, Perception, and Reality*, Ridgeview Publishing Company, Atascadero: 41-59.

Sellars, Wilfrid, 1963b, "Some Reflections on Language Games", in *Science, Perception, and Reality*, Ridgeview Publishing Company, Atascadero: 321-358.

Sellars, Wilfrid, 1963c, "Philosophy and the Scientific Image of Man", in *Science, Perception, and Reality*, Ridgeview Publishing Company, Atascadero: 1-40.

Sellars, Wilfrid, 1963d, "Empiricism and the Philosophy of Mind", in *Science, Perception, and Reality*, Ridgeview Publishing Company, Atascadero: 127-196.

Shapiro, Lionel, 2011, "Intentional Relations and the Sideways-On View: On McDowell's Critique of Sellars", in *European Journal of Philosophy*, 21,2: 300-319.

Tolman, Edward, 1932, *Purposive Behavior in Animals and Man*, The Century Company, New York.

Tolman, Edward, 1948, "Cognitive Maps in Rats and Men", in *Psychological Review*, 55,4: 189-208.

Weber, Andreas and Francisco Varela, 2002, "Life After Kant: Natural purposes and the autopoietic foundations of biological individuality", in *Phenomenology and the Cognitive Sciences*, 1: 97-125.

Wiener, Norbert, 1948, *Cybernetics, or Communication and Control in the Animal and in the Machine*, The MIT Press, Cambridge MA.

Wisdom, John, 1951, "The Hypothesis of Cybernetics", in *British Journal for the Philosophy of Science*, 2,5: 1-24.

Wittgenstein, Ludwig, 1974, *Tractatus Logico-Philosophicus*; Eng. tr. by D. Pears and B. McGuinness, Routledge, New York.

Wolfendale, Peter, 2019, "The Reformatting of *homo sapiens*", in *Angelaki*, 21,1: 55-66.

Emotion and affect in the space of reasons

Peter Olen

Abstract: Wilfrid Sellars's conception of "the space of reasons" makes critical assumptions about what constitutes persons and human uniqueness. Specifically, Sellars assumes that being human is defined through rationality. Although unique to Sellars, defining humans through rationality is an assumption not without its problems. I trace historical and contemporary issues with ignoring emotion and affect in our definition of persons and attempt to reconcile Sellars's commitment to behaviorism with a seeming conflict between rationality and emotion.

Keywords: Wilfrid Sellars; behaviorism; rationality; emotion; affect.

1. *Introduction*

Although known for a conception of persons defined through the space of reasons, Wilfrid Sellars was not the only mid-20th century philosopher to locate rationality as the defining characteristic of humanity. Grace de Laguna, Susanne Langer, and others denied the immediacy of awareness, endorsed versions of behaviorism, and argued that rationality was a (if not the) mark of the mental. Where these philosophers depart from Sellars, though, concerns what, exactly, picks out the essence of persons. For Sellars, the game of giving and asking for reasons is foundational for understanding the concept of personhood. To be a person is to be caught up in a web of practical and theoretical reasoning, moral and behavioral commitments, and "common sense" categories through which we experience the world. Characterizing persons without these categories fails to capture something crucial about what it is to be human. We just are the kinds of creatures that use these categories to experience and understand our world. And this characterization is not an individualistic account of persons; our role as members of a community (subject to the norms and conceptual inheritance of the group) is defining. The ability to see oneself through a uniquely human lens presupposes a sense in which we can see our-

philinq X, 1-2022, pp. 125-144
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.392
Submitted: May 2021
Accepted: October 2021

selves in the same light as others. To be a person is not to be a solitary reasoner, but to be a part of a community, one where members offer both collective and individual reasons for their behavior.

Failed accounts of human uniqueness are not exactly rare. Linguistic capacity, religious connotations, and morality have fallen to the wayside when characterizing ourselves as part of, yet somehow apart from, nature. Defining ourselves through reason seems like the last bastion for human uniqueness. Sellars's space of reasons is not a move against naturalism, yet there is a resistance to characterizing human practices as anything but rational. For creatures so strongly defined by our relationship to reason, we clearly do a poor job acting on, or for, reasons. Nonetheless, Sellars's account of persons requires a sense of rationality at both the individual and social level.

In what follows I explore neglected alternatives to Sellars's anchoring of persons to rationality. My contention is that the focus on reason or reason-giving as *the* defining characteristic of human experience has blinded us to alternative characterizations. This blind spot becomes most apparent when looking at notions of emotion, affect, collective reasoning, and group membership. Instead of characterizing the defining aspect of individuals and groups in terms of reason, the emotional and affective element should be seen as a key part of who we are. Sellars's depiction of persons, while leaving room for notions of emotion and affect, is largely silent on the issue. I am not suggesting that a Sellarsian account of persons is incapable of accounting for emotion or affect, but that current scholarship has ignored these aspects of experience in favor of rationality. More so, we have overlooked what was then-contemporary alternatives to Sellars's views (alternatives that also endorsed the importance of rationality, groups, and behaviorism when defining persons).

de Laguna functions as a historical alternative to Sellars because both philosophers were faced with the same cluster of intellectual developments, yet drew different conclusions from them: attempts to combine behaviorism, an emphasis on the social, and a naturalistic picture of persons led both philosophers down different roads. This historical connection supports my contention that Sellars's conception of persons is not an inevitable conclusion based on his endorsement of behaviorism, naturalism, and the social. de Laguna's shift in perspective might bring something valuable, yet overlooked, to a discussion of persons on both an individual and collective level.

2. *Human uniqueness and the space of reasons*

Although Sellars's conception of the space of reasons is primarily discussed in relation to his essay "Empiricism and the Philosophy of Mind",

his early distinction between human and non-human animals defines persons through rationality. Sellars is clear we should see a difference between non-human animal responses to environmental stimuli and the kind of rational choices made by human agents. Sellars draws a distinction between tied behavior ("learned responses to environmental stimuli") and free behavior that tracks the difference between humans and other creatures (Sellars 1949/2005: 122). While both human and non-human animals are conditioned to respond to environmental stimuli (either through brute habituation or social enculturation), only persons make use of symbol-laden activity, which constitutes the intellectual vision of our world. It is rule-regulated behavior that dominates Sellars's conception of persons: "To say that man is a rational animal, is to say that man is a creature not of *habits*, but of *rules*" (Sellars 1949/2005: 123).

There is a clear difference between emotional and rational characterization of animals. Emotion is reactive and involuntary. Insofar as I am emotional about something, I do not choose to be so (although I can choose to endorse my reactions to some experience and, with suitable practice, temper them). Reason, on the other hand, is at least partially intentional. In Sellars's philosophy this can be seen in the distinction between tied and free behavior. Though not his motivation for drawing the distinction, this distinction easily cleaves between emotional and rational characterizations of non-human animals and persons. Emotions are tied to stimuli and reactive.[1] Reasoning concerns my *choice* to think about things in a certain way (a choice I tend not to have when it comes to *feeling* angry or joyous). The characterization of persons as free and rational commits us to an intimate connection between both concepts. My ability to act because of reasons, instead of in accordance with them, signals a kind of humanity over nature.

Sellars's conception of persons in the manifest image is shot through with the idea that giving and asking for reasons is the hallmark of human practices. Insofar as we are concerned with articulating a conception of persons *from within* our "common sense"[2] image of the world, the categories and concepts through which we recognize others as rational are pivotal for human experi-

[1] Being reactive, though, does not mean being free of constraint by social and moral norms. Diaz and Reuter (2021) argue for the inherently normative character of emotions. Barrett (2018) argues viewing emotions as purely reactive misconstrues a sense of control we have over our emotions.

[2] As exhaustively noted by those writing on Sellars, the manifest image is not just an articulation of common sense, but a sophisticated, idealized image of the categories and practices through which we navigate the world. More helpfully juxtaposed against a purely scientific view of persons, the manifest image (in this instance) is being used as the sole perspective from which to discuss persons.

ence.[3] Given this, Sellars's analysis of practical engagement with the world does not really address any sense of emotional episodes, affective engagement, or embodiment.[4] It is not that emotion is completely absent; pains, desires, and feelings all make minor appearances, but only as ancillary considerations. One might think of sensations as covering emotions (as any emotional or affective element is felt, and said feeling is a sensation of something), but classifying emotions as sensations does not clarify the issue. Sellars's discussion of sensation is based on a perceptual model of the senses: sensations are caused by, or connected to, external stimuli in ways substantially different than emotions. While both visual sensations of objects and facial expressions involve becoming perceptually aware of external stimuli, the visual sensation of perceiving objects is drastically different from experiencing nostalgia or anxiety. We might see sensations as components of emotion, but they do not exhaust the concept.

The space of reasons need not be characterized as always containing explicit reason-giving practices. Sellars's characterization of humans *qua* the space of reasons turns on our *ability* to offer reasons for our behavior and to act *because* of them, but this does not mean all behavior involves an explicit practice of reason-giving. In the epistemic dimension of human practices, the role of reason is clear enough: insofar as we're making claims about what we know, what others know, what's true and what's false, prescribing a central role for reason-giving is a hallmark of epistemic practices.[5] Even if we take epistemic behavior as only a matter of reasoning, this does not disqualify emotional and affective considerations from our understanding of human cognition. Recent empirical research[6] has pointed out that the traditional, perceptual model of the relationship between reasoning, judgement, and emotion gets the order of

[3] Or at least one assumes so. Sellars say nothing about why we ought to privilege reason over all other options. The ease of which this assumption is accepted in Sellars's philosophy could be explained by the previous (and perhaps current) preoccupation with language. Given that Sellars's philosophy is mainly (if not exclusively) concerned with articulating an academic, and somewhat myopic, conception of persons, this doesn't mean other concepts or experiences do not play a major role in our lives.

[4] For example, Sellars's long discussion of moral and practical reasoning in the concluding chapter of Sellars 1967 says nothing about these issues. An extended and complex discussion of practical reason supports the very notion of what is it to be a person here, but little is said of emotional or affective states themselves.

[5] Rationality as the defining characteristic of persons is not only found in Sellars's work, but is pervasive throughout his intellectual descendants. John McDowell (1994), Robert Brandom (1994), and Joseph Rouse (2015) all emphasize the normative, rationally-constrained dimension of human cognition as defining for persons. There is little to no discussion of emotion and affect in their central works. Although de Laguna does not disagree with this view, she does provide a more robust role for emotion and affect in her work.

[6] For a summary of research surrounding affective realism see Barrett 2018, chapter four.

explanation wrong: it is not that reasoning and perception shape emotion and affect, but vice versa. Our emotional and affective states inherently shape our reasoning and judgment about issues. While there may be some conceptual room in Sellars's philosophy for this view, it seems unlikely to find a home.

Morality is traditionally the clearest place in human agency where emotion plays a central role. One finds Sellars's most robust (albeit still thin) discussion of emotion in his moral writings. Yet, Sellars relegates emotion to a causal or phenomenological role: emotion or affect boil down to aspects of motivation, they're part of the causal story behind our actions, or they're used in descriptions of moral experience. What is really doing the work, one imagines, is the "logic" behind one being caused to behave in specific ways. While emotion is a relevant consideration for Sellars (insofar as it plays some role in moving individuals to act), it is not the primary consideration when discussing morality.

Notice how the centrality of moral *reasoning* is already presupposed in Sellars's account of morality. This assumption makes sense if we are *starting* from the idea that moral reasoning is the primary concept in play for morality, but this serves as an unjustified starting point. There is a quasi-historical explanation available to us: Sellars frames morality as juxtaposed between intuitionist and emotivist conceptions of morality. These theories address emotion and affect, but fail to capture anything unique about morality through such concepts. This, in part, seems to be Sellars's motivation for thinking feeling isn't an adequate identifier for human morality. In addition, Sellars's adoption and modification of a Kantian conception of morality essentially guarantees the marginalization of emotion and affect. Take Sellars's discussion of obligation and motivation: there is a clear bifurcation between moral reasoning about obligation and any felt sense of responsibility. What is doing the work in his account of obligation is the "logical" structure of emotions and their role in rule-following (Sellars 1951). "Feeling obligated" is mentioned as a subject of empirical psychology, but this seems to badly misconstrue how ethical practices function within the manifest image. How emotion and affect shape moral experiences, guide and inform our actions, allow for and maintain relationships that constitute our ethical lives, and shape what it is to be contextually and rationally sensitive to morally salient considerations are all missing.

As another example, take Sellars's conception of materially valid inferences.[7] When thinking about moral experience, why start from the premise

[7] See Koons 2019, chapters 10 and 11 for an excellent exploration of this issue in relation to moral reasoning. Insofar as McDowell's solution to the "Humean problem" of moral motivation allows for a combination of emotion and reasoning, his solution is much closer to what I am proposing here. That being said, relegating emotion to a motivational role minimizes the far-reaching consequences of taking emotion and affect as human categories seriously.

that our actions arise as a result of some form of reasoning in isolation from emotional or affective elements? Insofar as I feel strongly compelled to intervene in a moral situation, I act. The phenomenology is instructive; even if my actions are guided by moral reasoning, it is unclear how, why, or when it comes into play. What moves me, in a literal and figurative sense, is the affective nature of being compelled to act. In experience, these aren't separate impulses, but simply what is found in human experience. It is, at best, a historical mistake to think reason-giving is separable from emotion and dominant within us. But even a discussion of emotion and affect in these cases is not simplistic. What caused me to intervene? The spectrum of emotional states and affective experiences is broad, perhaps guilt, perhaps anxiety or obligation. Infrequently, one imagines, it is moral reasoning or actions explicitly guided by principles that move us. Following Barrett, how the emotional and affective coloring of our experiences impacts reasoning is both absent in a reason-centric picture of persons and largely unexplored.

One objection might be that an emphasis on the role of emotions overlooks the rational nature of moral behavior. This is to reject my suggestions about emotion on the grounds that such seemingly immediate, non-deliberative action cannot be rational and, therefore, cannot be moral.[8] But this conflates acting for reasons with what picks out specifically moral behavior. As virtue theorist have argued since Aristotle, much moral behavior demands sensitivity to others and morally salient features of our experience, neither of which requires us to reason about issues. Frequently, doing the right thing is found in being sensitive to the right phenomena (which are not necessarily *considerations*). We can reconstruct this issues along reason-giving lines, but why should we?

The concept of persons present within this picture is, for the lack of better phrase, *hollow*. While Sellars's account of obligation may capture structural issues of morality, it fails to account for the lived, affective element that makes morality itself possible. We can, somewhat obviously, have abstract conceptions of morality that capture some dimension of moral reasoning. One's theory need not cover *all* aspects of morality. The problem is that a vision of morality that fails to include emotional elements risks creating a kind of fiction. To act as if one can explore the structural or logical aspects of morality without considering the role emotion plays in shaping our reasoning overlooks a crucial part of our experience (moral or otherwise). The conception of persons one gets out of this account perpetuates the idea that the moral or percep-

[8] Various moral psychologists have pushed back against the idea that rationality is the "mark of the mental". See Hindrinks and Sauer 2020 for summaries and arguments surrounding psychological rationalism.

tual dimensions of persons can be articulated without substantive reference to emotions and their impact on our conceptual capacities.

A Sellarsian response to this might be that affective states are simply folded into causal explanations of agency. *Of course* we feel something when reasoning about our obligations, but this felt state of awareness is both conceptually-laden and simply part of the causal story behind knowing. What is "really" doing the work is the reason-giving (though some felt state of awareness constitutes a necessary, though not sufficient condition for experience). Yet, this doesn't work as a defense of Sellars. Emotional states cannot simply be folded into a causal story without completely ignoring advances in conceptions of reasoning and emotion, as well as an important dimension of what marks human experience as uniquely human. Getting away from a traditional construal of emotions, one that not only places them in the backseat of knowing but ignores their constitutive role in reasoning itself, suggests that emotion cannot be put to the wayside.

Defining our conception of persons through the bifurcation of reason and emotion mistakenly drains all emotional and affective elements from morality. This can be seen clearly in Robert Binkley's discussion of how we ought to think of practical reasoning:

> The conception of logic as the science of reason needs perhaps a further comment. Reason exists only in reasoning, reasoning exists only in thinking, and only souls think, so logic in this sense is a kind of psychology. But it is a special kind of psychology for which some such special name as "rational psychology" had better be employed. This is to emphasize that while this logic is concerned with the forms of thought, it is concerned with them not as they reveal themselves to introspection, nor as they are manifested in behavior, nor even as they are related to physio-logical processes, but rather as they are reconstructed when we seek to represent our thought as rational. (Binkley 1965: 424)

Binkley's description of practical reasoning is instructive.[9] When held against Sellars's conception of moral and practical reasoning, Binkley's description provides an explicit methodological statement that is reflected in Sellars's philosophy. If we are constructing what moral and practical reason look like, not as actually practiced but as *imagined* as rational, then such an ideal-

[9] Why is Binkley's approach to practical reason important for understanding Sellars's view? As is clear in drafts of his "Imperatives, Intentions, and the Logic of 'Ought'", Sellars's discussion of practical reasoning – especially the sense in which it functions as a kind of rational reconstruction – is inspired by Binkley's work on the issue. See an early draft of Sellars's article at <https://digital.library.pitt.edu/islandora/object/pitt%3A31735062219211>

ized form of theorizing can be cut loose from emotional and affective bonds. This gives us an utterly rational depiction of persons, one that frames morality as an issue of acting from principles.[10] But notice, this is an *assumption* found in Sellars's philosophy. And this assumption is grounded in the notion that our primary engagement with morality is through reasoning. This doesn't preclude the inclusion of emotion and affect, but it denies the importance of such concepts. This is not done from an argumentative or evidential standpoint but is problematically assumed from the start – it is baked directly into Sellars's conception of persons.

The blind spot in Sellarsian accounts of persons just is the assumption that we can adequately describe or explain moral experience, meaningful experience, intentional thought, or any variety of uniquely human experiences as somehow devoid or absent of emotion and affect. Sellars's discussion of sensations is robust, but it is unlikely one could just extend that model to cover emotions and affect. The larger issue is that giving our emotional and affective experiences a more prominent place in characterizing human experience impacts our understanding of persons. There are more specific arguments that cast doubt on a traditional understanding of reason-giving as the primary explainer of specifically human behavior, but there is also the metaphysical issue of how we see ourselves. If constructivist accounts of emotion are correct, such as the ones found in Barrett's work, then we cannot make sense of an account of perception that is devoid of such supplementation. While Sellars leaves room for the inclusion of emotional and affective states in his picture of cognition, there is less room for a view of persons that is redefined by this change.

3. *Historical interlude*

The issue we are concerned with is not how reason became a defining feature of persons, per se; reason plays a central role throughout western philosophy's history. Our question is: *does this centrality create a blind spot in our characterization of persons?* Far from conclusive, there is nonetheless a historical narrative that begins to explain the move from 19th century science and philosophy to the position Sellars's found himself in during the mid-20th century. This historical setting matters because it provides *prima facie* evidence for the viability of conceptual and historical alternatives to Sellars's views, but also helps explain why there is a blind spot in Sellars's philosophy. A combination of developments in psychology (the move from introspective to behavioristic psychology) and the naturalization of social concepts (in the recognition of the

[10] See Sellars 1967: 203-205.

importance of collective groups) serve as historical guideposts. The argument here is that while behaviorism's mid-20th century prominence helps explain why Sellars's solely focuses on persons *qua* reasoners, this is by no means an automatic conclusion under behaviorism's momentary ascendancy.[11]

The behavioristic revolution in American psychology shifted emotions away from internal states *qua* intrinsically characterized episodes to a variety of thin and thick accounts of emotion. Some of the earliest behaviorist accounts of emotion can be found in simplistic descriptions of physiological or stimulus-response reactions (e.g. Watson 1919). More complex accounts of emotion, such as found in Edward Tolman (1923; 1932), take into consideration the "meaning" that might be entertained between felt, affective states and external stimuli, but they nonetheless depend on physical or physiological characterizations of emotional states and "unique 'directions' of behavior" to characterize different emotions (Tolman 1932: 268). Although these views entail different conceptions of emotion (albeit by seemingly minor degrees), they are unified against an introspective, intrinsically characterized conception of emotion and affect.[12] While affect is paid some lip service, it is still characterized in brute physical or behavioral terms.[13]

One way the emergence of behaviorism can be characterized is as a move from internal to external characterization of mental states. By "inner characterization" I mean something akin to introspective or "common sense" characterizations of human experience. Classically, insofar as I have an emotional experience, such experiences can be explained or characterized from my first-person standpoint. Physiological and behavioral terms could play some role in individuating emotions, but it is the *experience* of those emotions that defines them. A classic view (perhaps more indicative of philosophy than psychology) of inner episodes is the idea that thoughts and experiences *begin* in my immediate experience and (eventually) work outward to be expressed through language. Given the conceptual shift under behaviorist psychology, mental episodes characterized *from within* become intellectually suspect.

This change is perfectly encapsulated in Sellars's characterization of men-

[11] I am skeptical behaviorism's ascendancy can be classified as "momentary" for Sellars and Sellarsians. See Olen 2018 for an argument about the indispensability of behaviorism to Sellars's philosophy.

[12] For an explicit rejection of introspective accounts of emotion and affect, see Tolman 1932: 266-267.

[13] The history of psychological accounts of emotion is, of course, much more complex than this. One finds arguments about behavioral conditioning versus instinct, the order of physiological causes (i.e., do emotions cause physiological changes or do physiological changes cause emotions), and drastically different uses of "emotion", "feeling", and "affect". As to the latter point, see Dixon 2012 and Russell 2021 for explorations of these semantic differences.

tal episodes through his Rylean myth. Here, we find Sellars characterizing inner episodes through the use of external concepts. Specifically, the idea is that one can account for the introduction and "logic" of our private, inner episodes through public concepts used to describe language, overt behavior, and objects. In the Rylean myth, members of Sellars's mythical community are able to construct notions of inner episodes as modeled on the semantical categories applicable to overt behavior and language (Sellars 1956/2000: 267). This view of characterizing mental states turns on the idea that even states that seem to be intrinsically characterized can be given a public, intersubjective basis that is then internalized to play a reporting role for our experiences. Sellars is not rejecting the idea that we do introspect, but modifying such a notion to be consistent with behaviorism (Sellars 1956/2000: 264). Instead of starting from immediate, ostensibly unmediated, experiences and moving outward, Sellars's insight is that external concepts become internalized in order to classify and report our experiences. Behaviorism has no issue with introspection and internal episodes insofar as their occurrence is evidenced on behavioral grounds (Sellars 1962/1963: 22).

In light of psychology's conceptual shift, Sellars's modeling of inner episodes on external speech and behavior leads directly to a characterization of persons *qua* reasoners.[14] Insofar as we are modeling thought on speech, there is a unified structure and rationality behind the norm-governed use and intersubjective exchange of language. If thought is modeled on such an exchange, it stands to reason thought embodies the same structure found in natural language (or, at least, our explanation of thought invokes the same structural features). Given this line of reasoning, it is clear why Sellars and many of his contemporaries could be convinced of both the unified nature of practical reasoning and the second-class status of emotion.

Why think an intersubjective characterization of persons accurately depicts our experience of the world? This is a fairly complex question within Sellars's philosophy, as the accuracy of one's view will partially depend on the framework from within which we're discussing persons. There is no doubt Sellars thinks such concepts belong within the manifest image conception of "persons-in-the-world". Understanding behaviorism as a methodological restriction on concept formation helps explain any move away from internal characterizations of mental states.

Behavioristic commitments do not automatically entail the enshrining of

[14] With its emphasis on habit, one might think behaviorism is not an ideal candidate to embody persons *qua* reasoning. But it is the potent combination of behaviorism and Sellars's emphasis on language as *the* model for inner episodes that creates a conceptual blind spot.

reason or intentional thought over emotional states. de Laguna, for example, entertained both behaviorism and a complex conception of emotion and affect.[15] While earlier behavioristic treatments of emotion were somewhat simplistic, de Laguna (1919) offers a more complex explanation of emotional life rooted in behavioristic psychology. It is not that de Laguna initially tells a different story about emotion than most behaviorists; emotion and affect are still characterized in physiological and behavioral terms (de Laguna 1919: 418). What is different in de Laguna's case (and what functions as an entry point for a different behavioristic perspective) concerns the role emotion would play in characterizing human actions. Specifically, de Laguna makes an extended case for emotion and affect as a kind of unifying experience between human and non-human animals. Most forcefully seen in its role in collective integration and obligation (discussed below), emotion and affect drive our decision-making, motivation, and reasoning.

de Laguna does not ignore cognition or rationality as integral aspects of humanity. Despite emphasizing affect's role in cognition and obligation, de Laguna's view is still indicative of a traditional emphasis on rationality; she is clear that reason is the defining characteristic of persons (de Laguna 1927: 138). That being said, to juxtapose our choices between emotion and rationality as an exclusive disjunction presents us with a false dilemma. The point I am making by briefly mentioning de Laguna's position is that one can characterize persons as driven and characterized by emotion and affect, and *then* articulate a role for rationality.[16] Such an account does not contradict the use of external concepts for internal reporting roles nor an adherence to behaviorism. Even though rationality played an outsized role in characterizing persons, *it need not*.

Does this amount to a blind spot in Sellars's work? Despite the view of persons embodied in behaviorism, there are historical alternatives that create a more robust role for emotion in characterizations of persons. While Sellars's use of external speech and behavior as a model for inner episodes helps explain the move away from considering emotion as a dominant category for human experience, such explanations lose a bit of luster when considering alternative possibilities. The shift from introspective to behavioristic psychology

[15] Both de Laguna and Sellars insisted on the methodological character of behaviorism in strikingly similar terms. See Sellars 1956/2000: 263-266 and de Laguna 1927: 123-126.

[16] de Laguna's views are being presented as a "historical" alternative in the sense that her views were live options during Sellars's lifetime. This matters because one might think Sellars's views are, in some sense, an inevitability giving his cluster of commitments and historical epoch (e.g., to a Kantian sense of morality, to a form of behaviorism). This is simply not true, as de Laguna's philosophy exhibits.

does not require abandoning emotion in favor of reason when classifying persons. But once we decide that intrinsic characterization of mental states is out of play, combined with using overt behavior as the model for inner episodes, one begins to see why emotion plays such a minor role in Sellars's philosophy.

I say all of this not to offer a conclusive historical lens through which to see Sellars's conception of persons. Sellars, perhaps more than most, is resistant to being understood from just one perspective. Instead, I am arguing that the currents of intellectual history align in such a way that to overlook this conceptual development is all but impossible. The potent combination of behaviorism's movement from inner to outer characterization of concepts and the abstraction from actual human practices blinds us to alternatives.[17] This occurs not just on the level of individuals but can be seen even in broad conceptions of collective membership and obligation – both crucial notions for Sellars's conception of persons.

4. *Individuals and groups*

Sellars's conception of persons presupposes a normative framework of group membership. In Kantian fashion, Sellars's depiction of individuals rests upon a foundation of collective reasoning that supports moral, social, and epistemic practices. The principles and claims of practical reason, which both characterize human agency and account for our actions and obligations, arise out of the relationship between individuals and their communities. From Sellars's standpoint, insofar as I am interested in explaining the "logic" behind "I ought to do x", one cannot do so without some reference to "We ought to do x". Individual statements of obligation presuppose collective statements of obligation, while collective statements of obligation entail individual obligations.

What is important for our purposes is that Sellars defines group membership not just as thinking of oneself as part of a group, but in *reasoning as part of a group*.[18] Although Sellars does mention individual attitudes potentially being dependent on group attitudes, the primary way to think about being a member of a group comes from intentional thought: "I wish to emphasize that when the concept of a group is "internalized" as the concept of *us, it becomes a form of consciousness* and, in particular, a form of intending" (Sellars 1965: 203). This idea of a collective consciousness avoids being a naturalistically suspect group mind by internalizing the concept of a group. The metaphysical status of a collective,

[17] I've made similar arguments when it comes to Sellars's conception of language as well (see Olen 2016, especially chapters five and six).

[18] See Sellars 1962/1963: 39.

then, plays less of a problematic role in characterizing individual thought. Being part of a group means thinking and intending in a particular way.

As is the case with individual thought, we should question exactly why the focus is on reasoning, thinking, or intending when it comes to the relationship between individuals and their communities. Collective emotions, at least on the surface, are no more naturalistically suspect than collective intentions. There are a number of different ways in which we can think of collective emotions. There is the simple experience of being with others; a sense of togetherness that can foster collective feelings of comfort, relief, and safety. There is a sense of shared history or trauma that can dictate actions and reasoning about a variety of issues. All of these examples are largely emotional experiences with strong affective components. While reasoning or intending are a part of them, it would be odd to reduce traumatic experiences to "thinking of trauma as one of us".[19]

There are various ways of feeling like part of a community. Feeling pride in the idea that "we won the World Series" when the Tampa Bay Rays finally pull it off is a by-product of being part of a specific group. Red Sox's fans being angry – as a community – about their bitter loss can be embodied in individuals. But it isn't just the individual Red Sox fan being angry; it makes a substantial difference that *we* are angry about the Red Sox's loss. And, following Sellars's reasoning about disagreement within a group, one could paradoxically be part of a group, yet feel different from the group itself. The salient feature of collective emotions *qua* a shared, affective sense is (much like Sellars's idea of collective consciousness) the idea that *we* feel angry, happy, or prideful about a certain experience or event.

Representing these experiences as instances of reasoning or intending completely misunderstands the nature of those experiences. We *could* represent the process that leads to the exclamation "We won the pennant!" as the culmination of formally or materially valid inferences, as well as an intimate connection between intention and action, but that fails to capture both the emotional and affective dimensions of those experiences. More so, those experience simply would not be those experiences without the emotional and affective dimension. Reasoning alone cannot capture the salient features of our experiences as persons and members of a community.

One can see the same kind of difference within morality as well. Much like in the case of reasoning, there is all the difference in the world between whether I find something wrong and when we find something wrong. In the latter case, the *feeling* of general disagreement entails a number of different factors: it might make me more susceptible to agree or disagree with others,

[19] Helm (2014) does an excellent job of discussing the various models of collective emotion.

the fear of public shame might stifle my opinions or even cause me to speak up. Although reasoning can be an important component in these cases, it is unclear that reasoning is my primary reaction to these scenarios.

My point is not that collective intentionality fails to capture an important part of human cognition. Nor am I arguing that we should favor collective emotion over collective intentionality. In similar fashion to my discussion of reason and emotion, collective emotions should play a supplemental role in Sellars's philosophy and our understanding of persons. Sellars's views are not incompatible with a notion of collective emotion. Margaret Gilbert's approach to collective emotions (Gilbert 2014), for example, presupposes a normative framework of commitments and obligations in order to make sense of a non-summative account of collective emotions.[20] This framework depicts collective emotions as immersed within a network of social commitments and entitlements. So, there is a sense in which collective emotions fit easily into a Sellarsian framework. If we accept the idea collective emotions would be normatively-guided (i.e., that there is an important sense in which there is a right and wrong way to collectively feel in various instances), then the notion of collective emotion is no more problematic in Sellars's framework than the idea of collection intentionality.

One might argue that since Sellars is concerned with the rational articulation of human experience, emotion (what, as previously mentioned, is taken as a category for empirical psychology) need not play a role in their discussion of collective reasoning or intending. Aside from the fact that this functions exactly as the blind spot I discuss in sections II and III, this overlooks another salient feature of collective emotions: their *non-accidental* nature. To insist on a conception of collective emotion is not to claim some kind of accidental feature, such that you and I just happen to feel the same way about the Red Sox (disgusted by them; perhaps annoyed at their cheating ways). As a fan of the Tampa Bay Rays, I feel a *shared* sense of joy or pride in our crushing of the Red Sox. What marks such an experience is not any sort of collective reasoning about our victory, but *the emotional experience of being part of a group*. Just like reasoning, one expects collective emotions to share certain structural properties in common between individuals, too. Even if we think of collective emotions as found at the group level (i.e., as a kind of emotion that is somehow ascribable to the group itself instead of individuals), one imagines such concepts must be concretely instantiated in individuals at some point. But it would be experiencing or reacting as a member of the group that allows for this emotions to be manifested as collective in individual members of a community.

[20] Tuomela 2013 also contains a brief discussion of collective emotion.

Taking collective emotion as a supplemental, yet important, part of human experience does impact how we view persons. The problem, much like with Sellars's conception of reasoning, is that his initial conception of "group minds" and collective membership is simply devoid of emotional and affective content. We can keep the normative framework encapsulated in Sellars's talk of a "space of reasons", but we cannot do so in isolation from the role emotions play in cognition, behaviorism and our conception of persons.[21]

5. *Historical interlude, part two*

While de Laguna and Sellars wrote during different, yet overlapping, times (de Laguna being a contemporary of Sellars's father more than his peer), both philosophers found their way to the notion of collective intentions and the importance of the community through Emile Durkheim's work. While Stephen Turner and I have discussed Sellars's pathway to an idea of collective consciousness through Durkheim and Celestin Bougle, de Laguna's unique combination of behavioristic psychology and Durkheimian sociology has never been discussed.[22] Although not a mainstay of her work, there is a key passage that makes this connection:

> Behaviorism has interesting points of contact with the doctrine of the sociological school of Durkheim. The thinkers of that school are, to be sure, indifferent to any theoretical consideration of individual psychology, since it is a corner stone of their system that social phenomena are the subject-matter of a wholly independent science. But they are one with the behaviorists in insisting on the necessity of a thoroughly objective treatment of the phenomena in question. Social phenomena are, they admit, psychical and not physical or biological; but this does not imply that they are mental states or processes taking place in 'minds'. So far as '*representations collectives*' are open to scientific study, it is as objectively observable rites and instructions and formulated beliefs. Hence our own claim that the successful treatment of language depends on envisaging it as an objective phenomenon and in the light of its own objective relationships, instead of as a manifestation of inner mental states, is as much in accord with the spirit of Durkheim's sociology as it is with behaviorism.

[21] One assumes such a framework is necessary for any sense of emotion as non-accidental and collective. Stipulating certain connections and structural features of collective emotions only makes sense in the context of some form of normative framework. So, the argument throughout this essay is not that said framework should go away, but that – again – such a framework can only make sense when supplemented by a robust conception of individual and collective emotion.

[22] See Olen and Turner 2015.

> The influence of Durkheim's school on recent writers on linguistics is a significant symptom of a widespread trend of contemporary thought. The essential social character of language is more and more acknowledged and even insisted on in recent contributions to philological and psychological journals. What is lacking so far is the conception of the social *function* of speech. Speech continues to be referred to as the communication of ideas, which is still implicitly regarded as inner processes in individual minds. It would be far more in accord with Durkheim's general theory to regard the *function* of speech equally with the structure of language, as an objective social phenomenon. It is not the least merit of behaviorism that it provides a new view of the phenomena both of society and of the individual and of their interrelations. This does not mean 'reduction' of the one to the other, any more than the general program of behaviorism means a reduction of psychology to biology. (de Laguna 1927: 123-124)

de Laguna saw behaviorism as not only consistent, but adaptable to the notion of a collective consciousness. Both behaviorism and Durkheimian sociology insisted on an objective understanding of persons, one that flipped the traditional order of depending on internal explanations of our mental lives. Moving away from a view of language as an internal mechanism towards the idea of objective characterizations of language and thought is made possible through a combination of behaviorism's externalization of concepts and sociology's explications of the social basis of persons.

Combining behaviorism and the notion of a group mind might initially sound odd. Even in behaviorism's non-reductionist moods, the idea of a shared mind seems to push back against conceptual externalization. But the oddity of this combination makes more sense than it initially appears if one keeps in mind that de Laguna and Sellars 1) saw the chief virtues of behaviorism and social notions in their objective depiction of inner mental life and that 2) As explicitly stated in Sellars (and seemingly implicit in de Laguna's work), the notion of a group mind is simply not naturalistically-suspect under behaviorism's externalization of concept. Insofar as we internalize a notion of what "experiencing as one of us" is from publicly available resources, there is nothing naturalistically alarming.

The philosophical ramifications of these different paths sprouting from behaviorism and Durkheim are significant: despite the linguistic status of persons in a community, de Laguna argues emotion and affect shaped our reciprocal, felt sense of group membership and obligation (de Laguna 1927: 207). It is the *feeling* of compulsion, one potentially cashed out in behavioral terms, which gets the job done of naturalizing a seemingly ethereal entity like the group mind. Sellars, on the other hand, saw reason as amenable to naturalization and consistent with Durkheim's framing. Philosophically speaking, the

important question concerns whether affect or reason should play the central role in understanding a notion of collective agency is fundamentally the issue that falls out of this historical comparison. Sellars, for his part, says very little about affective states that live outside of sensory consciousness.

This conception of felt obligation morphs over de Laguna's career.[23] While she remained consistent about the importance of the affective dimension of our experiences, de Laguna shifted away from discussing these issues through the lens of behaviorism. Instead, one finds her offering broadly metaphysical speculations of human nature and enculturation. Here, one finds de Laguna offering similar arguments about the importance of affect, albeit under slightly different constraints. Affect shows up as an idealized concept under the guide of rationality (one imagines something similar to Gilbert's discussion of emotion), but the role remains the same: one of the most important aspects of group membership is feeling obligated (de Laguna 1963: 174-5). Regardless of these changes, the consistent message of the importance of affective senses of obligation play an important role in who we are. It is not just responding differential or thinking the right way about something, but also feeling "the right away" in response to others.

To be fair to Sellars, de Laguna does not have a substantially developed theory about the relationship between affect, the group, and the individual. Affect plays an intermediate role when moving from the newborn to enculturated person. While an affective basis may play as similar role in non-human animals as ourselves, we eventually behaviorally and linguistically respond to social cues that move beyond emotion and affect (de Laguna 1927: 214-215). This is not to say that emotion and affect drop out of the picture once we have become fully socialized creatures, but that we can be conditioned to respond to a variety of stimuli, in a variety of different ways, as mediated through the group. The alternative theory to Sellars, then, is not to replace reason with affect (or to ignore the way in which we become conditioned to respond to different perceptual and linguistic cues where emotion or affect may play a minimal role), but to acknowledge and explore the role emotion and affect play in shaping our socially cultivated form of human cognition.

More so, de Laguna's discussion of emotion and affect does not turn on a conception of collective emotions. While de Laguna is happy to discuss individual experiences of reasoning and emotion, her focus is general on the social basis of individual development. So, much like Sellars, a notion of collective emotion is consistent with de Laguna's understanding of emotion, behaviorism, and the social basis of our world.

[23] I have addressed this point in response to Joel Katzav's work on de Laguna in Olen $2$1.

6. *Conclusion*

In addition to highlighting a neglected alternative, my point has been to explore what we are overlooking when reason-giving takes center-stage. This is not to reject reasoning as *a* characteristic of persons in the manifest image but to show how focusing on reasoning alone overlooks essential aspects of human existence that define who we are. What these objections require is not an abonnement of the space of reasons or complex models of practical inference, but in true Sellarsian fashion, but supplementation from the additional categories of emotion and affect. I have not developed a theory about emotion or affect here, nor have I provided concrete evidence that de Laguna's position would have been a better choice. I have simply pointed out what is a meaningful and problematic oversight in Sellars's approach to persons.

Much of what Sellars is concerned with is the reconstruction of what *rational* human practices might look like. Such projects may find a useful home in both theoretical and practical concerns, but why think we should privilege this concern over phenomenology? Or why think that such an analysis must focus on reason to the detriment of all else? It is the presumption that reason plays a, if not *the*, central role in our experience that anchors such an assumption. More so, it is the conceptual space made by eliminating the connection between analysis and actual practices that makes room for a notion of reason devoid of emotion and affect.

But the empirical and conceptual points are only part of the picture. The fact that historical alternatives are present in mid-20 century philosophy, the fact that one could see a different path that would have offered a vital role for emotion to play in cognition and agency, is telling about our understanding of persons and their histories. Sections III and V, perhaps seeming disjointed with the conceptual arguments that run throughout the rest of the article, serve the purpose of showing how we can accept some of the best parts of Sellars's theories without giving in to a conception of persons that is hollow. Within their historical context, Sellars and de Laguna represent a set of shared premises – behavioristic commitments with a social twist – that led in different directions. These alternatives help show that we can keep the best parts of Sellars's theories while expanding his conception of persons to more fully account for the emotional elements of our experiences.

How we characterize persons does not carve nature at its joints, but it does substantially more than offer a socially constructed classification of individuals. Metaphysical concerns aside, the moral, political, and existential implications of this definition – while not fully realized within Sellars's philosophy – are pernicious. Holding a model of persons that focuses on theoretical and

practical reasoning misconstrues how we experience the world. Even if complex models of practical reason explain or characterize human behavior, it is unclear how such models graph onto our actual practices. The hope is that this paper functions more like a challenge than a full-blown argument. As Jeremy Koons and others have shown, Sellars's account of persons, practical reason, and ethics have much to contribute to the greater philosophical conversation – a contribution that is still overlooked. But Sellars's conception of persons needs supplementation; the emotional and affective elements at play in our experiences (or, better yet, that partially constitute those experiences) are simply absent in Sellars's philosophy, which leaves us with a diminished conception of persons.

Peter Olen
Lake Sumter State College
olenp@lssc.edu

References

Barrett, Lisa, 2018, *How Emotions are Made: The Secret Life of the Brain*, Houghton Mifflin Harcourt Publishing Company, New York.

Binkley, Robert, 1965, "A Theory of Practical Reason", in *The Philosophical Review*, 74: 423-448.

Brandom, Robert, 1994, *Making It Explicit: Reasoning, Representing, and Discursive Commitment*, Harvard University Press, Cambridge.

de Laguna, Grace, 1919, "Emotion and Perception from the Behaviorist Standpoint", in *The Psychological Review*, 26: 409-227.

de Laguna, Grace, 1927, *Speech: Its Function and Development*, Indiana University Press, Bloomington.

de Laguna, Grace, 1963, "The Person", in *The Review of Metaphysics*, 17: 171-186.

Diaz, Rodrigo and Reuter, Kevin, 2021, "Feeling the Right Way: Normative Influences on People's Use of Emotion Concepts", in *Mind and Language*, 36: 451-470.

Dixon, Thomas, 2012, "Emotion": The History of a Keyword in Crisis, in *Emotion Review*, 4: 338-344.

Gilbert, Margaret, 2014, "How We Feel", in von Scheve, Christian and Salemla, Mikko, eds. *Collective Emotions*, Oxford University Press, Oxford: 18-30.

Helm, Bennett, 2014, "Emotional Communities of Respect" in von Scheve, Christian and Salemla, Mikko, eds. *Collective Emotions*, Oxford University Press, Oxford: 47-59.

Hindriks, Frank and Sauer, Hanno, 2020, "The Mark of the Moral: Beyond the Sentimentalist Turn", in *Philosophical Psychology*, 33: 569-591.

Koons, Jeremy, 2019, *The Ethics of Wilfrid Sellars*, Routledge, New York.

McDowell, John, 1994, *Mind and World*, Harvard University Press, Cambridge.

Olen, Peter, 2016, *Wilfrid Sellars and the Foundations of Normativity*, Palgrave Macmillan, London.

Olen, Peter, 2018, "The Varieties and Origins of Wilfrid Sellars's Behaviorism" in Corti, Luca and Nunziante, Antonio, eds. *Sellars and the History of Modern Philosophy*, Routledge, New York: 178-196.

Olen, Peter, forthcoming, "Perspectivism and Behaviorism: A Response to Katzav", in *Australasian Philosophical Review*.

Olen, Peter and Turner, Stephen, 2015, "Durkheim, Sellars, and the Origins of Collective Intentionality", in *British Journal for the History of Philosophy*, 23: 954-975.

Russell, James, 2021, "Psychological Construction of Episodes Called Emotions", in *History of Psychology*, 24: 116-120.

Sellars, Wilfrid, 1949/2005, "Language, Rules and Behavior", in *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*, ed. by J. Sicha, Ridgeview Publishing Company, Atascadero: 117-134.

Sellars, Wilfrid, 1951, "Obligation and Motivation", in *Philosophical Studies*, 2: 21-25.

Sellars, Wilfrid, 1956/2000, "Empiricism and the Philosophy of Mind", in deVries, Willem and Triplett, Tim, eds. *Knowledge, Mind, and the Given: Reading Wilfrid Sellars's "Empiricism and the Philosophy of Mind"*, Hackett Publishing Company, Indianapolis: 205-276.

Sellars, Wilfrid, 1962/1963, "Philosophy and the Scientific Image of Man" in *Science, Perception and Reality*, Ridgeview Publishing Company, Atascadero: 1-40.

Sellars, Wilfrid, 1965, "Imperatives, Intentions, and the Logic of 'Ought'", in Castaneda, Hector-Neri and Nakhnikian, George, eds. *Morality and the Language of Conduct*, Wayne State University Press, Detroit: 159-218.

Sellars, Wilfrid, 1967, *Science and Metaphysics: Variations on Kantian Themes*, Ridgeview Publishing Company, Atascadero.

Tolman, Edward, 1923, "A Behavioristic Account of the Emotions", in *Psychological Review*, 30: 217-227.

Tolman, Edward, 1932, *Purposive Behavior in Animals and Men*, University of California Press, Berkeley.

Tuomela, Raimo, 2013, *Social Ontology: Collective Intentionality and Group Agents*, Oxford University Press, Oxford.

Watson, John, 1919, "A Schematic Outline of the Emotions", in *Psychological Review*, 26: 165-196.

Does philosophical knowledge presuppose a moral attitude? A discussion of Max Scheler's metaphilosophical thesis

Íngrid Vendrell Ferran

Abstract: This paper explores Max Scheler's metaphilosophical views. In particular, the paper seeks to reconstruct and assess Scheler's thesis according to which philosophical knowledge presupposes a moral attitude which he describes as an "act of upsurge" on the part of the whole person of the philosopher toward the essential, an act which cannot be found in either the natural worldview or the sciences. After motivating the topic in the introduction (section 1), the paper explores how Scheler approaches the question about the nature of philosophy by focusing on the type of person of the philosopher (section 2). It then examines Scheler's claim according to which philosophy is fundamentally distinct from the sciences (section 3), before exploring the moral attitude of the philosopher by examining three of its conditions: love, self-humbling, and self-mastery (section 4). The paper presents some challenges and objections against Scheler's metaphilosophical thesis. In particular, critiques of its metaphysical implications and of the view of science implicit in it are provided (section 5). Finally, it is also argued that the thesis contains a grain of truth and as such a moderate interpretation of it could be defended (section 6). The main findings are summarized in the conclusion (section 7).

Keywords: Max Scheler; phenomenological metaphilosophy; scientific attitude; essence; Platonism; metaphysics; love; humility; virtue.

1. *Introduction: Max Scheler's metaphilosophical thesis*

Since the second half of the 20th century, metaphilosophy has developed into a field of philosophical inquiry concerned with questions regarding the aims, nature, methods, and values of philosophy itself (Cath 2011). Though it is not clear how metaphilosophy can be demarcated from other philosophical subdisciplines, there is general agreement that it approaches the old questions of philosophy – such as What is philosophy? What is its purpose? How should we philosophize? Is philosophy a science? – to reflect upon the activity of philosophizing itself. One of its aims consists of searching for the different answers that have been provided to such questions in the course of the history of philosophy. In this vein, there has been a growing interest in investigating the

philinq X, 1-2022, pp. 145-168
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.391
Submitted: July 2021
Accepted: October 2021

metaphilosophical views defended by the most influential Western currents of thought of the last century. For instance, the entry "Metaphilosophy", written by Joll (2017) for *The Internet Encyclopedia of Philosophy* examines metaphilosophy in pragmatism, analytic philosophy, and phenomenology. Moreover, publications on the metaphilosophical views of prominent authors in the history of philosophy have proliferated. To mention but a few, in recent years studies on Kant's, Nietzsche's or Wittgenstein's metaphilosophies have appeared. The present paper is conceived against this background and aims at enlarging today's metaphilosophical debate by exploring Max Scheler's (1874–1928) metaphilosophical views. In particular, the paper reconstructs and assesses an intriguing idea which can be found in Scheler's article "On the Essence of Philosophy and the Moral Condition of Philosophical Knowledge" (Vom Wesen der Philosophie und die moralischen Bedingung des philosophischen Erkennens, 1917). According to this idea, to which Scheler explicitly refers as a "thesis" (1954: 79; Eng. tr. 2010a: 85), philosophical knowledge presupposes a moral attitude which he, in quite Platonic terms, describes as an "act of upsurge" (Akt des Aufschwungs) of the whole person of the philosopher toward the essential, an act which cannot be found in either the natural worldview or the sciences.[1]

What are the motivations for writing a paper on Scheler's metaphilosophy? To begin with, the existing literature on phenomenological views in metaphilosophy is devoted mainly to Husserl, to existentialist phenomenologists such as Sartre and Merleau-Ponty, and to later authors influenced by the phenomenological movement such as Derrida.[2] A focus on these major figures has led to a neglect of the metaphilosophical views of other, lesser known authors such as Scheler whose original insights are also worth acknowledging if we want to obtain a more complete picture of metaphilosophical views within phenomenology. Second, metaphilosophy has been dominated mainly by the question of the method or methods of philosophy,[3] often in combination with questions regarding the differences and similarities between philosophy and natural science. By contrast, other intriguing questions – such as the one posed by Scheler

[1] In this paper, I will indicate the page numbers of the original German as well as of the English translation. Though for quotations I will use the English translation, I have employed the original German for the reconstruction of Scheler's thought.

[2] A good example of this focus on these authors is the above-mentioned entry "Metaphilosophy" written by Joll (2017). In their book, Overgaard *et al.* examine explicitly analytic philosophy and continental philosophy, the latter of which is often associated with phenomenology (2013).

[3] Cath (2011) argues that in the analytic tradition, metaphilosophy has focused mainly on questions of method. The same diagnosis can be made for metaphilosophy in the phenomenological tradition in which questions regarding the phenomenological method have been central too.

regarding a hypothetical moral attitude adopted by the philosopher in order to attain philosophical knowledge – have received scant attention. Finally, the very idea that philosophical knowledge requires a moral attitude is intriguing in itself. Whether or not we agree with Scheler's specific understanding of this moral attitude, the question as such prompts us to think about the conditions of philosophizing and about the type of person we are as philosophers.

With the aim to present and discuss what I call here Scheler's metaphilosophical thesis, I will adopt two methodological strategies. First, I will reconstruct Scheler's thesis connecting his thoughts contained in the 1917 article mentioned above with other works of the same period. As noted by Schloßberger, Scheler is "a system thinker" (2020: 72). Therefore, the understanding of any specific systematic topic cannot be approached in isolation because in Scheler's work, ethical, epistemological, and metaphysical issues are intertwined in multiple ways. In particular, as I will demonstrate below, Scheler's metaphilosophical thesis is intimately linked to his philosophy of mind (and, specifically, his philosophy of affectivity), his ethics, and his metaphysics. Second, in order to assess his thesis, I will approach his work by focusing on three systematic questions: 1) the question about the nature of philosophy; 2) the question about the relation between philosophy and the sciences; and 3) the question about the moral conditions required to attain philosophical knowledge. Systematizing Scheler's thoughts around these three issues will enable us to better understand the arguments that sustain his thesis and will also pave the way to assess its plausibility and potential for current research.

The structure of the paper will be as follows. After motivating the topic, the paper will first explore how Scheler approaches the question about the nature of philosophy by focusing on the philosopher as a type of person (section 2). It then examines Scheler's claim according to which philosophy is fundamentally distinct from the sciences (section 3), before exploring the moral attitude of the philosopher by examining three of its conditions: love, self-humbling, and self-mastery (section 4). The paper presents some challenges and objections against Scheler's metaphilosophical thesis. In particular, critiques of its metaphysical implications and of the view of science implicit in it are provided (section 5). Finally, it is also argued that the thesis contains a grain of truth and as such a moderate interpretation of it could be defended (section 6). The main findings are summarized in the conclusion (section 7).

2. *Defining philosophy by the philosopher's spiritual attitude*

When Scheler approaches the question about the nature of philosophy in the aforementioned text of 1917, he notes that compared to the empirical sci-

ences (which he designates the "positive sciences"), philosophy cannot easily answer the question about what philosophy is. However, at the same time, Scheler considers the question about the nature of philosophy to be in fact a question of "philosophy's self-knowledge through philosophy" (1954: 63; Eng. tr. 2010a: 70). That is, the question "what is philosophy?" is constitutive of philosophy itself.

In order to reconstruct his position, let me begin by sketching what in my view are, for Scheler, two problematic strategies that try to answer this question and define philosophy. The first such strategy is what I call the *conceptual approach*. He states that when we try to define what chemistry, physics or psychology are, we can always resort to a philosophical explanation of the main concepts of these disciplines such as matter, energy, consciousness or life. By contrast, philosophy, which, according to Scheler, constitutes itself through the question about its nature, cannot be defined by resorting to a philosophical explanation of its main concepts. Any attempt to define philosophy employing this approach would entail a circular argument because to know whether or not a content is philosophical presupposes that we already have an idea about what philosophy is and what is its object. In addition, in philosophy, we do not have a fixed doctrine or a system to which we can resort in order to answer this question. A second strategy which he finds inappropriate is what I call the *historical approach* which involves consulting the history of the discipline itself. This would presuppose that we already have an idea about the essence of what different authors at different times have called "philosophy". Though these strategies might be useful to find out the nature of other disciplines, they are useless when it comes to elaborating a definition of philosophy.

According to Scheler, the question about the nature of philosophy has to take as its point of departure the *autonomy* of philosophy. That philosophy is autonomous means that it cannot presuppose as true knowledge of its history, knowledge of the natural sciences, knowledge of the natural worldview or knowledge obtained through revelation. Philosophical knowledge is, for Scheler, the most unconditional form of knowledge. Thus, any attempt to define philosophy by resorting to knowledge of other disciplines would imply a form of traditionalism, scientificism, fideism, or dogmatism.

Having rejected conceptual and historical approaches and determined the autonomous character of philosophical knowledge, Scheler suggests answering the question about the nature of philosophy by looking at the philosopher as a "type of person" (Persontypus) (1954: 64; Eng. tr. 2010a: 70). I want to call this the *person's type approach*. This approach is valid only for autonomous disciplines (in the sense mentioned above). According to Scheler, it is false to think that it is easier to delimit a "subject matter" or a "task" than to indicate

or to recognize the "type of person" who possesses competences for such matters and tasks. Thus, although there is a general skepticism toward defining art as what the true artist makes, religion as what the true saint experiences, and philosophy as the relation to things exhibited by the true philosopher, Scheler thinks that at least as a heuristic tool we can determine a subject matter by examining the type of person who possesses competences for it.

That said, Scheler observes that when we decide whether a person such as Plato, Aristotle or Descartes is a "true" philosopher, there must be an idea that guides us in taking this decision. For Scheler, this guiding idea upon which it seems to be a certain implicit agreement but whose content is not always clear to us is not an empirical concept. In fact, this idea is about a fundamental spiritual attitude (geistige Grundhaltung) toward things that is characteristic of the personality of the philosopher. For Scheler, this idea which is still hidden is what guides us when we decide whether or not a person is a philosopher.

In Scheler's view, by focusing on the type of person of the philosopher and her spiritual fundamental attitude toward things, we can find out the nature of the object of philosophy itself. Against the widespread view that philosophy does not have its own objects and that it studies the same objects as the sciences but from another point of view, Scheler argues that philosophy does indeed have its own subject area (Sachgebiet), world of objects (Gegenstandswelt) or world of facts (Welt von Tatsachen) (1954: 65-66: Eng. tr. 2010a: 71-72). Although these facts exist independently of us, they are accessible only through a specific spiritual attitude. Thus, to determine what is the object of philosophy and what we can know about it, it is first necessary to examine the philosophical spiritual attitude that we have in mind when we claim that a person X is a philosopher.

What is this spiritual attitude? For Scheler, ancient philosophers discovered that the object of philosophy is in a particular realm of *being* (Reiche des Seins) (1954: 67; Eng. tr. 2010a: 73). They discovered that contact with this realm of being is linked to a specific act (Aktus) in which the whole personality is involved and which is missing from the natural attitude. This act was for the ancient philosophers an act of moral nature since it presupposes that we overcome a hindrance common to the natural attitude that prevents us from entering into contact with the being of philosophy. In Scheler's view, this act was already described by Plato as a "movement of the soul's wings" and as an "act of upsurge" (Akt des Aufschwungs) (1954: 67; Eng. tr. 2010a: 73) of the whole person toward the essential of all possible things. The essential is not a special object that exists beside the empirical objects; rather it is the essential in all possible things. This *dynamis* at the core of the person was described as an "eros", i.e., as a tendency or movement of all incomplete beings toward a complete being.

Though Scheler does not adopt the entire platonic doctrine, he takes from Plato two ideas regarding the fundamental spiritual attitude necessary for philosophy. First, Scheler adopts from Plato the idea that it is an act of the core of the person that puts us in contact with the object of philosophy. This act cannot be found in the natural worldview (natürliche Weltanschauung) nor in the empirical sciences which are founded on such a view. The second idea that Scheler takes from Plato is that this act is founded on a specific kind of love. For Scheler, philosophy underlies "a love-determined movement of the inmost personal Self of a finite being toward participation in the essential reality of all possibles" (1954: 68; Eng. tr. 2010a: 74). This means that a human being of the type of the "philosopher" is a human being who is able to adopt this attitude toward the world.

Yet, for Scheler, the spiritual philosophical attitude is not exhausted in these two moments. To them must be added a third moment according to which philosophy is knowing and the philosopher is one who knows. It should be stressed here that Scheler's notion of knowledge is more far-reaching than the usual understanding of knowledge in terms of justified true belief. More specifically, Scheler calls knowledge the participation of a finite human being in essential reality. This view of knowledge is not only defended in this text, but permeates Scheler's philosophical production (see, for instance, "Erkenntnis und Arbeit" (Knowledge and work) 1960: 227).[4] Accordingly, Scheler understands philosophy as an a priori self-evident intuition of essences and of essential interrelations of being (1954: 98; Eng. tr. 2010a: 104). These essences are not just the result of applying the phenomenological method as in Husserl; rather they constitute a realm of facts which can be disclosed if we adopt a specific spiritual attitude toward them.[5] They are not subordinated to empirical reality or to thought; they transcend us.[6] This is the world of objects of philosophical knowledge which, according to him, can be approached by focusing on the philosopher's spiritual attitude.

Before highlighting the main features of this spiritual attitude in comparison to the attitudes we have in the natural worldview and in the sciences, and before determining its specific moral conditions, let us step back for a moment

[4] In "Erkenntnis und Arbeit", Scheler distinguishes between different forms of knowledge in more detail (283). For an analysis of the relation between these different forms of knowledge and love, see Vendrell Ferran 2015.

[5] Scheler here takes for granted the idea that we have a priori knowledge about different domains of reality. For the idea of a priori knowledge in realist phenomenology, see Smith 1997, and for Scheler's specific development of this idea, see Kelly 2012.

[6] For the idea that phenomenology discloses a realm of facts in Scheler, see Meyer 1987: 21 and Mohr 2012: 229.

and consider Scheler's approach from the perspective of current metaphilosophy. As the above reconstruction of Scheler's approach shows, his proposal is to define philosophy by examining the person-type of the philosopher. Yet, as we have seen, in Scheler's work, his approach to the question of the nature of philosophy is intimately linked to what can be called an "essentialist" position.[7] His essentialism can be stated in at least two different respects. First, he takes as a point of departure the idea that all activities called "philosophy" have something in common, i.e., there is something that can be called the "essence" (Wesen) of philosophy. This idea is reflected clearly in the German title of the essay which contains the word "Wesen", i.e., "essence" (in the English version, "Wesen" is translated as "Nature"). As a result, though he proposes, as a heuristic tool, that we approach the realm of objects of philosophy by focusing on the type of person of the philosopher, he is not advocating for a deflationist position according to which philosophy is just what people called philosophers do. It is not his aim to answer the question about the nature of philosophy empirically, observing what people called philosophers have de facto in common. Rather, he approaches the question by looking at the idea that guides us when we claim that someone is a philosopher. This idea is for him a spiritual attitude toward the world. It is worth noting that Scheler's approach is both descriptive and normative. He not only describes the spiritual attitude that guides us when we call someone a philosopher, but also regards this attitude as the attitude that philosophers must exhibit.

Second, Scheler is also an "essentialist" regarding the objects of philosophical scrutiny. This issue is important here because in general philosophers do not agree about what exactly they study. Unlike history and psychology (and other disciplines such as biology, zoology, astronomy, physics, etc.) which have clearly delimited research fields, the question about what philosophy studies has been the topic of significant controversy. At one extreme there are those who do not know what exactly is the object of philosophy, while at the opposite extreme there are those who claim that the question is irrelevant, and in between them lie a wide and variegated range of positions about what is the object of philosophical scrutiny (for an overview, see Overgaard *et al.* 2013: 2-ff.). Against this backdrop, Scheler claims that philosophy has its own world of objects to which he refers as the essential of all possible things. Indeed, for Scheler, philosophy is the intuition of essences. As we will see in the next section, it is precisely because philosophy deals with essences that it requires a moral spiritual attitude which is absent from the natural worldview and science. Yet, these essences, as Scheler himself states in the third "Preface" to

[7] For essentialist and deflationist positions in metaphilosophy, see Overgaard *et al.* 2013.

his *Formalism* book, are not like Platonic ideas. Rather, as he writes: "I reject, in principle and at the very threshold of philosophy, a heavenly realm of ideas and values that is 'independent' of the essence and execution of living spiritual acts, independent not only with regard to man and human consciousness but also with regard to the essence and execution of a living spirit in general" (1973a: XXX). In short, Scheler's essences are not completely independent of the beings who are capable of grasping them.

3. *The natural worldview and the distinction between the philosopher's and the scientist's attitudes*

As described above, in order to attain philosophical knowledge, we have to adopt a moral attitude and overcome certain barriers that we encounter in the natural attitude that hinder us from entering into contact with the realm of objects of philosophy. Since for Scheler, science too implies a change in attitude regarding our natural worldview wherein we are immersed in our dealings with the environment, in which we take our environment to be "the world", and in which we are focused on the region of being relative to the sphere of vitality, a comparison between the philosophical and the scientific attitudes is required. Indeed, as we will see, the changes in attitude presupposed by philosophy on the one hand, and the sciences on the other, differ in substantial respects. In other words, for Scheler, philosophy is fundamentally distinct from the sciences, a term that he deliberately employs in the plural.

To start with, some historical remarks are in order. What we today call "science" emerged gradually out of philosophy during the Renaissance and early Modernity. During this time, issues which were regarded as typically philosophical became the subject matter of the new disciplines of physics, chemistry, biology, etc. This development robustly questioned the function of philosophy with respect to the sciences and left the door open to different interpretations of the relation between the two. It is in the context of these different interpretations that Scheler discusses in his text four positions: philosophy as the queen of the sciences; philosophy as the servant of the sciences; philosophy as itself a science; and finally, philosophy as distinct from science, which is the position with which he aligns himself.

Scheler argues first that the ancient idea that philosophy is autonomous and as such distinct from science – an idea which was best exemplified in the view of philosophy as "the queen of the sciences" (*regina scientiarum*) – developed into the opposite view according to which philosophy is "the handmaiden of the sciences" (*ancilla scientiarum*). According to this latter view, the main function of philosophy is to unify the results of the sciences (as is the case

for positivism) or to examine their conditions and methods (as it is the case for the scientific philosophy). However, for Scheler this view and the development that led to it are based on an "overthrow of all order of values" (Umsturz aller Wertordnung) (1954: 74; Eng. tr. 2010a: 79). This idea of an inversion of values that we find in this 1917 text was already employed by Scheler some years earlier in his essay on *Ressentiment* written in 1912 (2010b), in which he argues that when the desired goals cannot be achieved, a distorted apprehension of values and their hierarchy ensues. Genuine values and their bearers are then degraded, and this devaluation leads to genuine values being replaced by illusory ones. The distortion of the heart consists precisely in this inversion of values and this replacement. According to my reading, though not mentioning this emotional attitude in the 1917 text, Scheler is appealing to the structure of Ressentiment when he refers to an inversion of values. In particular, he is resorting to his critique of modernity, and its moral subjectivism, egalitarianism, and the negation of high values and how modernity involves an inversion of values according to which the useful appears to be the most important of all values.

That said, the view that philosophy is a form of science is also unacceptable to Scheler. In this respect, his views differ substantially from Husserl's well-known claim defended in his *Logos* article "Philosophy as Rigorous Science" (1911) that philosophy should be a rigorous science. Husserl seems to have maintained this view at least until the *Crisis of the European Sciences and Transcendental Phenomenology* (1936). While a thorough comparison of Husserl's and Scheler's views would deserve an article of its own, in what follows let me sketch the main points of convergence and divergence between the two authors.

Scheler himself exhibits some similarities with Husserl. First, like Husserl, Scheler distinguishes self-evident knowledge of essences (Wesenserkenntnis) from positive knowledge (Realerkenntnis). In addition, Scheler also argues that while positive knowledge remains in the sphere of probability, "philosophy is self-evident knowledge of essences" (1954: 75; Eng. tr. 2010a: 80). Third, like Husserl, Scheler also distinguishes philosophy from the deductive sciences of "ideal objects" (mathematics, logics, and theory of numbers).[8]

[8] Scheler states that Husserl expresses higher esteem for the phenomenology of the act and the phenomenology of the psychic than for the phenomenology of distinct regions of being (e.g., the phenomenology of nature) and confesses to find this preference unwarranted. This observation is important in the light of Husserl's turn toward transcendental philosophy and his interest in the constitutive activities of consciousness which is in contrast to realist phenomenology which was more interested in the analysis of the regions of being (for both directions within the phenomenological movement, see Geiger 1933).

Yet, despite these points of agreement, there are profound differences between the two authors. These differences are terminological as well as substantial.[9] The first disagreement concerns the use of the term "science". According to Scheler, Husserl employs this term with two different meanings: one for philosophy as self-evident knowledge of essences (what Husserl calls "rigorous science"); the other for the positive formal sciences of ideal objects and the empirical sciences. In contrast to Husserl, Scheler prefers to reserve the term for the latter meaning, employing for the former the term "philosophy" (Scheler 1954: 75; 2010a: 81). Therefore, though the German word for science, "Wissenschaft", can be employed to refer to the natural as well as the human sciences, Scheler prefers to use the term for the natural (e.g., biology) and formal sciences (e.g., mathematics), and to concede philosophy a place of its own. As he puts it, philosophy is nothing but philosophy and, as such, it possesses its own idea of strictness and of its discipline, and philosophy need not be ruled by any ideal notion of scientific discipline.

A second relevant difference between Scheler and Husserl concerns the words "Weltanschauung" and "Weltanschauungsphilosophie". The term "Weltanschauung", which in English translations is often quoted in German and means something like "vision of life" or "worldview", is employed by Scheler (who on this point follows von Humboldt) to refer to the forms of apprehending and envisaging the world which prevail at a given time over a given area as well as the structure of given intuitions and values of various social units such as peoples, nations, and civilizations. One finds these "Weltanschauungen" in the syntax of languages, religions, and ethos. The "Weltanschauungsphilosophie" is for Scheler the philosophy which is a natural constant for the human being. By contrast, Husserl employs the term "Weltanschauungsphilosophie" for what Scheler calls "scientific philosophy", i.e., the attempt made within the frame of positivism to develop a metaphysics or "Weltanschauung" that takes the results of science as its point of departure. Scheler agrees with Husserl that "scientific philosophy" is absurd. However, he does not agree with Husserl in calling the "scientific philosophy" a "Weltanschauungsphilosophie" because for Scheler, "Weltanschauungen" "evolve and grow"; they are not thought up by scholars (1954: 77; Eng. tr. 2010a: 83).

Scheler also agrees with Husserl that philosophy itself cannot be a "Weltanschauung", but at most only a doctrine of "Weltanschauungen" (Weltanschauungslehre). He agrees with the idea that creating a doctrine of particular

[9] Some of the disagreements between Husserl and Scheler mentioned here have already been noted by Mohr 2012. Here I complement and extend his work providing an analysis of Scheler's main motivations for the claim that philosophy is distinct from the sciences.

"Weltanschauungen" (such as the Christian or Indian worldview) is not the main task of philosophy. However, for Scheler, there is a philosophy of "natural Weltanschauungen". This "Weltanschauungslehre" would be, with the help of phenomenology, able to assess the cognitive value of "Weltanschauungen". It would also show that the structures of the prevailing "Weltanschauungen" occasion and influence the structure, character, and level of science effective in a society at a given time and would show that any variation in the structure of science is preceded by an analogous variation in the "Weltanschauung". Thus, while Husserl tends to give the positive sciences a greater factual independence from the changing "Weltanschauungen", Scheler considers the sciences to be dependent on prevailing "Weltanschauungen". In fact, for Scheler, the structures of science, their factual systems of fundamental concepts and principles always take place within the given structures of a "Weltanschauung".

As mentioned above, a complete understanding of the analogies and differences between Scheler and Husserl regarding the relation between philosophy and science/the sciences, despite its historical interest, is beyond the scope of this paper. Here my aim is rather to shed light on the reasons why Scheler, unlike Husserl, regards philosophy as fundamentally distinct from the sciences. To this end, I will focus on what I take to be the main four criteria motivating Scheler's view of a fundamental distinction between philosophy and the sciences.

3.1. Objects and their mode of givenness

The first criterion concerns the objects studied by philosophy and science. These disciplines' objects are, for Scheler, of a different kind. According to Scheler, philosophy has its own objects which cannot be reduced to the objects which occupy us in the natural attitude or in the sciences. This view was stated already in his article "The Theory of Three Facts" written between 1911 and 1912, where he distinguishes three kinds of facts: facts given in the common-sense experience of the natural worldview; facts which are studied in the natural sciences; and phenomenological facts which are revealed in the eidetic intuition (Scheler 1973b: 215).[10]

Moreover, science works with an "artificial" worldview and deals with states of affairs gained through observation (1973b: 226). The objects' different forms of existence result in different forms of givenness in the sciences. For instance, natural sciences require an extraverted attitude, while psychology requires an introverted attitude (1954: 84; Eng. tr. 2010a: 90). By contrast, these variegated forms of givenness cannot be found in philosophy. The mode of givenness in philosophy is intuition.

[10] For an analysis of this point, see Schutz 1957: 307.

Furthermore, unlike philosophy, which is intuition of essences and essential connections, science moves in the sphere of contingency. Even if science seeks universal laws of nature and presupposes knowledge of essences, it is unable to provide such knowledge.

3.2. Abilities

The second criterion which, according to my reading, motivates Scheler's distinction between philosophy and the sciences concerns the abilities involved in both activities. By virtue of the nature of their respective objects (numbers, geometrical figures, plants, animals, etc.), the sciences require the exercise of partial abilities of the human being. What Scheler means by this is that some sciences require more observation while others more reasoning; some sciences are more deductive while others are more intuitive, etc. While in the sciences only a part of the person of the scientist is involved, philosophy requires the involvement of the whole human being. Even when the philosopher approaches a specific question, in philosophy, it is the whole person of the philosopher who is involved in this activity (Scheler 1954: 84; 2010a: 90). An important consequence of these differences regarding the involvement of our abilities in the realization of a task is that for Scheler, while philosophy is one, the sciences are many.

3.3. Values, goods, and aims

Let us turn to what I consider to be a third criterion behind Scheler's distinction. This criterion concerns the values, goods, and aims of philosophy on the one hand, and the sciences on the other. As a follow up of the second criterion, those disciplines that are linked to certain types of values and types of goods such as art, law, politics, etc. require a one-sided application and exercise of emotional functions. For instance, art requires a sense for qualities, legal sciences require a feeling for what is fair, just, etc. By contrast, in philosophy, even when it deals with very specific problems, it is the whole human being who philosophizes and not only her reason or her sensibility. For Scheler, this thesis, which can also be found in Plato, is not a psychological thesis, but an ontical one (1954: 84-85; Eng. tr. 2010a: 90).

Furthermore, and this is a central point for Scheler, the scientist is moved by practical aims. Her goals are the "control and modification" of the surrounding world (Beherrschabarkeit und Veränderlichkeit) (1954: 91; Eng. tr. 2010a: 97). Unlike the scientist, the philosopher controls and modifies only insofar as this enables her to enter into contact with a sphere of absolute being. Thus, though both exhibit the same attitude of "self-mastery", this activity is employed with different aims.

3.4. Moral or non-moral nature of the activities

Finally, philosophy and the sciences differ regarding their moral or non-moral nature. Scheler argues that though both philosophy and the sciences presuppose an attitude which differs from the natural attitude, it is philosophy but not science that requires a moral attitude that makes us transcend the natural worldview and puts us in contact with essences. Philosophy requires an act of upsurge that puts us in contact with a different realm of being. In this context, Scheler defends the radical view that the goal of the human being in the upsurge is to create a unity of her being with the being of the essential and, in so doing, to transcend herself. As I will elaborate below, this moral attitude is only possible because the philosopher is moved by love.

Though science too requires an attitude which differs from the natural worldview, the attitude required by science is not moral because the scientist does not aim at entering into contact with essences.[11] Moreover, though the contents of science are different from those we target in the natural attitude, in its formal structure science remains based in the natural attitude (1954: 89; 2010a: 94). In fact, the scientist must love knowledge, but unlike the philosopher, she must not love the being of things.

These fundamental differences between philosophy and the sciences are based on the idea that philosophy deals with a region of objects which is beyond the reach not only of the natural attitude but also of the empirical and formal sciences. A view of this sort can be regarded as a form of "Platonism".[12] As described by Overgaard *et al.*, "Platonism" is the view that philosophy deals with a "deep" and intangible part of reality beyond the reach of the sciences (2013: 32). In Plato's allegory of the cave formulated in his *Republic*, while the empirical sciences study the world of shadows, the philosopher accesses the region of the intelligible and is able to contemplate the true forms of the beauty, the good, etc. Though Scheler rejects the idea of an independent realm of ideas and values, for him philosophy and the sciences deal with different regions of objects, access to which requires different attitudes. Moreover, the objects of philosophy, which Scheler regards as a realm of the essential, seem to be placed at a more elevated region than the objects of the sciences. Precisely for this reason, for him, the idea of philosophy as "ancilla scientiarum" is the expression of an inversion of values.

[11] Note that Scheler's view is compatible with the idea that scientific praxis can be guided by moral principles.

[12] Scheler's Platonism differs from Husserl's Platonism in one crucial respect. While Husserl argues that philosophy should be a rigorous science, Scheler, like Plato, places the sciences at a different level than philosophy. For a comparative analysis of Husserl's and Plato's views of the relation between philosophy and the sciences, see Arnold 2018, 35 and ff.

4. *Characterizing the philosopher's spiritual attitude in terms of moral conditions*

4.1. The Primacy of Love Over Practical and Theoretical Reason

This section delves deeper into Scheler's idea about a moral condition for philosophical knowledge. In what sense does the spiritual attitude of the philosopher have a moral character?

Scheler warns us not to conflate his claim that philosophy requires a moral condition with Kant's and Fichte's idea of a "primacy of practical over theoretical reason" (Primat der praktischen Vernunft vor der theoretischen) (1954: 78; Eng. tr. 2010a: 84). Scheler's moral condition should not be interpreted in terms of a practical attitude toward the world because this attitude is for Scheler linked to the natural attitude and as such it can be deceptive. In this respect, Scheler regards his account as being closer (though not entirely identical) to the views defended by the ancient philosophers for whom a moral spiritual attitude (the act of upsurge) is necessary to obtain philosophical knowledge, i.e., for entering into contact with the realm of being which is the object of philosophy. Yet, unlike the ancients for whom the theoretical life was regarded as more valuable than the practical one, Scheler does not argue for a "primacy of theoretical reason". In fact, what Scheler defends is what he himself calls a "primacy of love".

Scheler's argument for the claim that philosophical knowledge presupposes a moral condition presumes a certain familiarity with theses defended in his other works of that period. In what follows, resorting to his epistemology, philosophy of affectivity, ethics, and metaphysics, I will shed light on the three main tenets of his argument. To this end, I will proceed in three steps.

First, Scheler starts by showing that we must become aware of the motives of self-deception that may misguide us in the apprehension of value. In this vein, he claims that "authority and education" (Autorität und Erziehung) are required to achieve an intuition of value (and the will and action founded on it) (1954: 79; Eng. tr. 2010a: 85). To acquire the ability to see what is right and good, one must overcome the motives for self-deception which precede the intuition of values and which are responsible for deceiving us about values and for making us blind to them. The motives for self-deception that Scheler mentions in the paper on the essence of philosophy are forms of life which consist mainly in objective bad will and action turned habitual for us. Scheler highlights these self-deceptive forms in order to demonstrate that the apprehension of value requires a moral condition. Yet, these are for Scheler by no means the only sources of self-deception. Though he does not develop the is-

sue here, in his essay on the "Idols of Self-knowledge" written in 1911 (1973c), Scheler makes clear that we have a natural tendency to take what is given in our environment as if it were our own. This involves other forms of self-deception than those mentioned in the 1917 text. For instance, we might think and even have the emotional illusion that we are sad because we are attending a funeral, without in fact being sad about it. For Scheler, these self-deceptive tendencies need to be clear to us so that we can apprehend values without falling victim to self-deceptive maneuvers. As the example illustrates, there are forms of self-deception about values that are formed in a pre-volitional sphere, i.e., before actions, and that are capable of conditioning the will, action, and emotional experience.[13]

One might ask here why Scheler speaks of a moral condition for the apprehension of values when he in fact wants to demonstrate that what requires a moral condition is philosophical knowledge, a knowledge that he interprets as the cognition of being. Indeed, Scheler makes clear that the cognition of values presupposes a moral condition, but he wants to show that the cognition of being presupposes a moral condition too. To this end, in the next step of his argument, he shows that there is an essential connection between the cognition of value and the cognition of being. For him, no value-free being can be the object of a perception, memory, expectation, thought or judgment. In fact, there is a logical (not temporal) order of foundation according to which an apprehension of value precedes the apprehension of things. A being which is value-free or value-indifferent is for Scheler the result of an artificial abstraction. In fact, as Scheler stated in his work *The Nature of Sympathy* (2008), first published in 1913, we are emotionally involved with things and only come to regard them as value-free via an abstraction.[14] In his view, the consciousness of values precedes the consciousness of things. Scheler also endorses this strong thesis in his *Formalism*, where he clearly states: "A value precedes its object: it is the first 'messenger' of its particular nature" (1973a: 18). Again, the apprehension of values (a phenomenon that he calls "value-ception") "*precedes* all representational acts according to an essential law of its origins" (1973a: 201). In short, for Scheler, values are given to us prior to the entities which are their bearers.

So far, according to Scheler's argument, the givenness of values is prior to the givenness of being (though values are only attributes of being) and since the givenness of values presupposes a moral condition, then the givenness of

[13] For an analysis of the biosemiotics aspects of Scheler's thought, see Cusinato 2018.

[14] This idea reappears in the 1917 essay when, for instance, he argues that the Pitagorean treated numbers as if they were deities before analyzing them.

being also (indirectly) presupposes it. Yet, how does Scheler end up defending here a "primacy of love"? To understand this next step of his argument, it is important to know that for Scheler values are given in affective acts. This thesis is stated in several works of the period from 1910 to 1920. Scheler regards love as the most primary of all acts and the one that makes us open to the world of values and being. Thus, for him, before something is known, it must be first loved or hated: "Everywhere the 'amateur' precedes the 'savant'" (1954: 81; Eng. tr. 2010a: 86). Or as he puts it in "Ordo amoris": "Man, before he is an *ens cogitans* or an *ens volens*, is an *ens amans*" (1973d: 110-111). Therefore, given that for Scheler values are given in affective acts and that among all affective acts those responsible for disclosing values are acts of love, he argues here for a "primacy of love". Now we can understand how Scheler's "primacy of love" is distinct from Kant's and Fichte's thesis of a "primacy of practical over theoretical reason" as well as from the ancient thesis of a "primacy of reason over practical life" (although, as he himself acknowledges, his view is closer to the latter than to the former because, like Plato, he argues that there is a moral condition of philosophical knowledge).

Let me stress here that for Scheler love is not an emotion but an attitude of openness of the human being toward the world. As he puts it in his *Formalism*, love is a primordial act or basic attitude that forms the core of our personality and enables our access to the world of values: "Love and hate are acts in which the value-realm accessible to the feeling of a being […] is either extended or narrowed […]" (1973a: 261). While emotions are, in Scheler's view, responses to values (for instance, fear is a response to an object presented as dangerous), love is not a response but a form of being open toward the other and her positive higher values. In this respect, Scheler's love has a disclosive nature.

4.2. The Moral Act of Upsurge: Love, Self-humbling, and Self-mastery

Having explained Scheler's main argument for the moral condition of philosophical knowledge, let's turn to his particular understanding of the "moral act of upsurge" which puts us in contact with the realm of being that is the object of philosophical knowledge. As mentioned in section 3, this act presupposes that we overcome certain barriers and inhibitions which are typical of the natural attitude whereby we are immersed in our environment and consider things only insofar as they are relevant for our vital purposes. In this act of moral upsurge, Scheler identifies three acts which must work together in unitary interaction (Scheler 1954: 89-90; Eng. tr. 2010a: 95). From these three acts, only the first is of a positive nature, while the second and third are of a negative character because in them we have to refrain from tendencies inherent to our being. For Scheler, *only these three moral acts* enable the human be-

ing to participate in the sphere of absolute being and to overcome the natural egocentrism (natürlichen Egozentrismus), vitalism, and anthropomorphism (1954: 90; 2010a: 95) which are typical of the natural worldview. Each one of these acts can be regarded as a moral condition for philosophical knowledge:

1. Love of the whole spiritual person toward absolute value and being. This love breaks the source of the relativity of being and leads us in the direction of absolute being. As stated above, for Scheler, love is a form of openness toward higher values and as such it has a disclosive function. Moreover, it is more basic than cognition and will. He regards this love as the motor of the entire complex of acts.
2. Self-humbling (Verdemütigung) of the natural self and I. According to Scheler this act breaks the natural pride (Stolz) and leads us from the contingent existence of something toward its essence.
3. Self-mastery (Selbstbeherrschung) and objectification of the instinctual impulses of life which are given as "bodily" and experienced as founded on the body, and which condition the natural sensory perception. This act overcomes the natural concupiscence and leads us from an inadequate perception of objects to adequate intuitive knowledge. For Scheler, self-mastery is the only activity shared by the philosopher and the scientist. Yet, while the scientist employs it to control and change the world, for the philosopher self-mastery enables the act of self-humbling.

It is important to note, as Schutz has done (1957: 306), that Scheler is speaking of a spiritual moral attitude and not of a technique or a method. The "true" philosopher is not a technician, but a type of person who is able of love, to become humbled, and to achieve self-mastery in order to transcend the natural attitude with the aim to discover a realm of philosophical facts.

Moreover, let me stress that for Scheler, these three acts do not just have an epistemic function, but also serve to open us toward the sphere of absolute being. Therefore, Scheler finishes his investigation by stating that inquiry into the nature of philosophy must begin with the question of the order of fundamental self-evident insights. And, for him, the first self-evident insight is that "there is something", or put otherwise "there is not nothing". The second insight is that there is an absolute entity, i.e., an entity which is not dependent on other entities. The third insight is that every possible entity must necessarily possess an essence and an existence. Therefore, the philosopher's spiritual attitude is not just an epistemic attitude that enables us to attain philosophical knowledge. Rather, it is an attitude of openness toward the sphere of absolute being.

In the light of these last reflections, it can be said that Scheler's metaphilosophical thesis is in fact a metaphysical thesis about how finite being comes to participate in essential reality (as stated in section 1, he defines knowledge

precisely in terms of participation). This result might seem radical, but it is interesting in two key respects. First, if we take as a point of departure the idea already noted by some Scheler scholars such as Meyer (1987) and, more recently, Mohr (2012) that Scheler's philosophy is imbued by metaphysical assumptions, it is not surprising that his metaphilosophy is in fact a version of his metaphysics. Second, this result is particularly interesting in the light of a strong thesis presented by Geldsetzer in the 1970s according to which metaphilosophy is in fact another name for metaphysics (Geldsetzer 1974: 255). Though I think we should not hurry to adopt Geldsetzer's radical view, the results of this section should invite us to rethink the relation between both philosophical subdisciplines.

5. *Challenges and objections against Scheler's metaphilosophical thesis*

So far, I have reconstructed and discussed Scheler's metaphilosophical thesis according to which philosophical knowledge presupposes a moral attitude. More specifically, I have shown that for Scheler love, self-humbling, and self-mastery are the three moral preconditions that enable us to overcome certain tendencies inherent to the natural attitude so that we can enter into contact with a world of essences which is the object of philosophical knowledge. In this section, my aim is to assess the plausibility of this thesis and to evaluate its potential for current research. To this end, I will present a series of challenges and possible objections against his view.

The first series of objections against Scheler's thesis concerns his essentialism, i.e., the view that all activities we call philosophy have something in common which he describes as an act of upsurge in the direction of a realm of essences. Scheler's essentialism is problematic because it leaves aside many other conceptions of philosophy present in Western and non-Western thought which do not work with the idea that philosophy is a form of intuition or that it has to do with essences. Moreover, the notion of "essence" central to phenomenology has been subjected to strong criticism from other currents of thought. As shown by Overgaard (2010: 902), Ryle already attacked Husserl for investigating "super-objects" called "essences". This attack is the result of a tendency to interpret the notion of essence as independent of the beings able to apprehend them. However, as I mentioned above, Scheler explicitly rejects the view of a "heavenly realm of ideas".

Second, when Scheler describes the attitude necessary to attain philosophical knowledge, he seems to describe phenomenology itself. In fact, in "Phenomenology and the Theory of Cognition" written between 1913 and 1914, Scheler argues that one of the central aspects of the group of thinkers called

phenomenologists is that they are inspired by a common attitude (Einstellung). As he puts it: "phenomenology is neither the name of a new science nor a substitute for the word 'philosophy'; it is the name of an attitude of spiritual seeing in which one can see (er-schauen) or experience (er-leben) something which otherwise remains hidden, namely, a realm of facts of a particular kind. I say 'attitude,' not 'method'" (1973e: 137). While a method is a procedure of "thinking about facts", in phenomenology it is a matter of "new facts themselves" and of a procedure of "seeing". Though Scheler insists that phenomenology is not a new name for philosophy, in fact, his understanding of philosophy is very close to what he calls phenomenology. A problematic consequence of this view is that what he regards as "true" philosophers are either phenomenologists who adopt the phenomenological attitude or authors working close to the standards of the phenomenological attitude. In this view, philosophers working with other understandings of philosophy would not be "true" philosophers.

Third, Scheler's views of philosophy and the sciences can be challenged in the following three respects. To begin, as we have seen, he defends a version of Platonism and considers philosophy as being occupied with a deeper and more elevated dimension of reality than science (Overgaard *et al*. 2013: 33). This view does not interpret science and philosophy at the same level, but rather sees them in a hierarchical relation in which philosophy stays above science. For Scheler, this hierarchy is based on the kind of objects studied by each of these disciplines: essences are deeper than states of affairs gained via observation. However, if we take a different perspective, for instance, the point of view of the consequences, Scheler's thesis can be strongly challenged because science's capacity to improve our life, at least from this vantage-point, is "higher" than philosophy's capacity. It is science, not philosophy, that cures disease, takes us to Mars, enables us to explore micro and macrocosmic worlds, etc.

Moreover, Scheler's thesis presupposes that philosophy is an "armchair" discipline distant from empirical concerns. However, philosophical knowledge can be informed, motivated, corrected or molded by looking into the results of the empirical disciplines and, conversely, philosophical views might influence, correct, mold or motivate empirical research.

Furthermore, Scheler's idea that philosophy and the sciences share only the attitude that he calls "self-mastery" is controversial. In my view, both philosophy and science require attitudes akin to Scheler's "love" and "self-humbling". In the current debate on virtue epistemology, open-mindedness and intellectual humility are often mentioned as important virtues in philosophy and science. My thought here is that a parallelism between love and open-mindedness, on the one hand, and self-humbling and humility, on the other, could be traced in order to show that the scientist is not only moved by a wish

to control and manipulate the environment, but she also must exhibit other epistemic virtues and some of them come close to the moral activities which according to Scheler are exhibited by the philosopher. Though this parallelism is obviously imperfect because Scheler does not speak of virtues but of moral attitudes and acts,[15] my idea against Scheler is that both the philosopher and the scientist must show different competences in order to excel in their task.

6. *Reformulating Scheler's thesis in terms of virtue epistemology*

These challenges and objections are important because they demonstrate that Scheler's metaphilosophical thesis cannot be accepted without simultaneously committing ourselves to problematic metaphysical views. Should then the thesis that philosophical knowledge presupposes a moral attitude be rejected? In my view, Scheler's thesis entails a kernel of truth insofar as it claims that true philosophers exhibit certain competences that make them able to attain philosophical knowledge. In line with this thought, in this section, I suggest a moderate interpretation of Scheler's thesis in what I think are less controversial terms. More precisely, I will gesture toward a reformulation of his thesis in terms of current virtue epistemology. In so doing, I will leave aside the metaphysical idea that the moral attitudes required to attain philosophical knowledge aim at putting us in contact with a realm of absolute being.

In this regard, I want to end this paper with a proposal to explore the extent to which Scheler's thesis that love, self-humbling, and self-mastery are moral preconditions of philosophical knowledge can be made fruitful for current research as a thesis about epistemic virtues necessary for philosophical knowledge. Clearly, this proposal moves us away from Scheler, but in my view it nevertheless opens new and intriguing paths of research about the abilities we must exercise in order to attain philosophical knowledge.

My thought here is that Scheler's moral conditions of love, self-humbling, and self-mastery can be transposed into more contemporary terms by looking at the debate on virtue epistemology. Since an elaboration of all these conditions is beyond the context of this paper, I will consider here just one of them: self-humbling. Consider in more detail the concept of intellectual humility. In virtue epistemology, intellectual humility is described as an epistemic character trait whose exercise might lead us to obtain epistemic value. In contemporary accounts the term is used in different meanings. For instance, Whitcomb

[15] Though the phenomenon of "self-humbling" is not described by Scheler in terms of a virtue, in his essay on "The Rehabilitation of Virtue" in which he examines humility, he explicitly employs this term (Scheler 2005: 24).

et al. focus on three meanings of the term (2017: 512-516) while Snow distinguishes eight meanings (Snow 2019: 178-195). Some of these accounts interpret intellectual humility in terms of a disposition but, interestingly, in her account, Tanesini has defended the idea that humility is a cluster of attitudes. As she puts it: "it is a cluster of strong attitudes (…) directed toward one's cognitive make-up and its components, together with the cognitive and affective states that constitute their contents or bases, which serve knowledge and value-expressive functions" (Tanesini 2018: 399). Yet Tanesini interprets this attitude as configured partly by modesty, a respect for one's intellectual strengths, and an acceptance of one's limitations. For Scheler, humility is distinct from modesty (Bescheidenheit). First of all, modesty but not humility is associated with shame. In addition, while modesty is a social attitude which can be understood only with reference to other human beings, humility is a personal and individual act. In spite of these differences, my thought here is that Scheler's concept of self-humbling could be transposed into more contemporary concepts such as that of intellectual humility.

Second, it should be investigated whether or not the three conditions mentioned by Scheler are exhaustive or whether the list needs to be extended. In my view, though love, self-humbling, and self-mastery are important, these are not the only attitudes required by the philosopher. She must also be sensible, honest, courageous, purposeful, and determined, to mention but a few.

Third, it should be examined whether these activities come in different proportions when we philosophize and when we do science.

Finally, it should be examined to what extent the moral attitudes mentioned by Scheler can be translated in terms of virtues, and whether these virtues are moral but also fulfill an epistemic function, or whether these are epistemic virtues outside the moral domain.

In developing these different transpositions, we would leave behind Scheler's wider project and its strong commitment to metaphysical views. Yet, I also think that this is possible because Scheler's project is also a pedagogical project. Moreover, my proposal here should be an attractive one insofar as I am suggesting that we can reinterpret and develop Scheler's thesis in terms of current virtue epistemology yet without committing ourselves to his metaphysics: namely the idea that the philosopher must exhibit certain moral attitudes in order to excel in the pursuit of her task.

7. *Concluding remarks*

In this paper, I have explored Scheler's metaphilosophical thesis according to which philosophy presupposes a moral attitude. From the reconstruction

and discussion of Scheler's thesis elaborated from sections 2 to 4, three main claims can be extracted. First, Scheler is an essentialist regarding the nature of philosophy and the objects of philosophical knowledge. More specifically, he regards philosophy as intuition of essences. Second, unlike Husserl for whom philosophy should be a rigorous science, Scheler defends the view that philosophy is not only fundamentally different from the sciences, but it also deals with objects which are at a more elevated level. Third, Scheler argues that philosophical knowledge requires the moral conditions of love, self-humbling, and self-mastery. Yet this thesis is intimately linked to his philosophy of affectivity, ethics, and metaphysics: the three mentioned acts enable finite beings to participate in essential reality. Section 5 evaluated the plausibility of Scheler's metaphilosophical thesis and argued that the thesis cannot be accepted without simultaneously accepting controversial metaphysical claims. In section 6, I called for a reformulation of the thesis, transposing his idea of moral attitudes in terms of virtues required by the philosopher in order to excel in her task, and suggested to examine whether these virtues differ in kind, proportion, and goals from those exhibited by the scientist.[16]

Íngrid Vendrell Ferran
Goethe-Universität Frankfurt am Main
vendrellferran@em.uni-frankfurt.de

References

Arnold, Thomas, 2018, *Phänomenologie als Platonismus. Zu den Platonischen Wesensmomenten der Philosophie Edmund Husserls*, De Gruyter, Berlin/Boston.

Cath, Yuri, 2011, "Metaphilosophy", in Pritchard, Duncan, ed. *Oxford Bibliographies Online: Philosophy*, Oxford University Press, New York.

Cusinato, Guido, 2018, *Biosemiotica e psicopatologia dell'ordo amoris. In dialogo con Max Scheler*, FrancoAngeli, Milano.

Geiger, Moritz, 1933, "Alexander Pfänders methodische Stellung", in Heller, Ernst and Löw, Friedrich, eds. *Neue Münchener Philosophische Abhandlungen*, Johann Ambrosius Barth, Leipzig: 1-16.

Geldsetzer, Lutz, 1974, "Metaphilosophie als Metaphysik: Zur Hermeneutik der Bestimmung der Philosophie", in *Zeitschrift für allgemeine Wissenschaftstheorie / Journal for General Philosophy of Science*, 5, 2: 247-255.

[16] I am grateful to Danilo Manca and Giacomo Turbanti for inviting me to submit a paper for their special issue. I am indebted to Matthias Schloßberger and Alexander N. Wendt for their insightful comments on an early draft and to Simon Mussell for proofreading the paper. This paper has also strongly benefited from the comments provided by two anonymous reviewers.

Joll, Nicholas, 2017, "Metaphilosophy", in *Internet Encyclopedia of Philosophy* <https://iep.utm.edu/con-meta/#H4>.

Kelly, Eugene, 2012, "Max Scheler", in Luft Sebastian and Overgaard Søren, eds. *The Routledge Companion to Phenomenology*, Routledge, Oxon-New York: 40-49.

Meyer, Herbert H., 1987, "Max Scheler's Understanding of the Phenomenological Method", in *International Studies in Philosophy*, 19, 1: 21-31.

Mohr, Eric J., 2012, "Phenomenological Intuition and the Problem of Philosophy as Method and Science: Scheler and Husserl", in *Symposium*, 16, 2: 218-234.

Overgaard, Søren, Gilbert, Paul, and Burnwood Stephen, 2013, *An Introduction to Metaphilosophy*, Cambridge University Press, New York.

Overgaard, Søren, 2010, "Royaumont Revisited", in *British Journal for the History of Philosophy*, 18, 5: 899-924.

Scheler, Max, 1954, "Vom Wesen der Philosophie und der moralischen Bedingung des philosophischen Erkennens", in *Gesammelte Werke* 5, Francke Verlag, Bern-München: 61-99.

Scheler, Max, 1960, "Erkenntnis und Arbeit", in *Gesammelte Werke* 8, Francke Verlag, Bern/München.

Scheler, Max, 1973a, *Formalism in Ethics and Non-formal Ethics of Values*, Evanston, Northwestern University Studies in Phenomenology and Existential Philosophy.

Scheler, Max, 1973b, "The Theory of Three Facts", in *Selected Philosophical Essays*, Evanston, Northwestern University Studies in Phenomenology and Existential Philosophy: 202-287.

Scheler, Max,1973c, "The Idols of Self-Knowledge", in *Selected Philosophical Essays*, Evanston, Northwestern University Studies in Phenomenology and Existential Philosophy: 3-97.

Scheler, Max, 1973d, "Ordo amoris", in *Selected Philosophical Essays*, Evanston, Northwestern University Studies in Phenomenology and Existential Philosophy: 98-135.

Scheler, Max, 1973e, "Phenomenology and the Theory of Cognition", in *Selected Philosophical Essays*, Evanston, Northwestern University Studies in Phenomenology and Existential Philosophy: 136-201.

Scheler, Max, 2005, "On the Rehabilitation of Virtue", in *American Catholic Philosophical Quarterly*, 79, 1: 21-37.

Scheler, Max, 2008, *The Nature of Sympathy*, Transaction, New Brunswick.

Scheler, Max, 2010a, "The Nature of Philosophy and the Moral Preconditions of Philosophical Knowledge", in *On the Eternal in Man*, Transaction, New Brunswick: 67-104.

Scheler, Max, 2010b, *Ressentiment*, Marquette University Press, Milwaukee.

Schloßberger, Mathias, 2020, "Max Scheler", in Szanto, Thomas and Landweer, Hilge, eds. *The Routledge Handbook of Phenomenology of Emotion*, Routledge, London: 72-86.

Schutz, Alfred, 1957, "Max Scheler's Epistemology and Ethics", in *The Review of*

Metaphysics 11, 2: 304-314.

Smith, Barry, 1997, "Realistic Phenomenology", in Embree, Lester, ed. *Encyclopedia of Phenomenology*, Kluwer Academic Publishers, Dordrecht: 586-590.

Snow, Nancy E., 2019, "Intellectual Humility", in Battaly Heather, ed., *The Routledge Handbook of Virtue Epistemology*, Routledge, New York: 178-195.

Tanesini, Alessandra, 2018, "Intellectual Humility as Attitude", in *Phenomenology and Phenomenological Research*, 96, 2: 399-420.

Vendrell Ferran, Íngrid, 2015, "Pragmatismuskritik und Wertbegriff bei Max Scheler", in Gabel, Michael and Müller, Matthias, eds. *Erkennen – Handeln – Bewähren. Phänomenologie und Pragmatismus*, Traugott-Bautz, Nordhausen: 120-134.

Whitcomb, Dennis, Battaly, Heather, Baehr, Jason, and Howard-Snyder, Daniel, 2017, "Intellectual Humility: Owning Our Limitations", in *Philosophy and Phenomenological Research*, 94, 3: 509-539.

Metaphilosophy of the Life-world
Sellars, Husserl, and the quantum image of nature

Danilo Manca[1]

Abstract: The aim of this article is to assess whether the notion of "life-world" could be helpful for a philosophical theory that assigns a primacy to the scientific view of the world when it comes to establish what exists. I set out to integrate the concept of "life-world" as understood in Husserl's late phenomenology with the point of view defended by Sellars in *Philosophy and the Scientific Image of Man in the World*. In what follows, I will consider the image of nature proposed by the standard "Copenhagen" version of quantum physics. This will allow me to challenge Sellars's assumptions that reality cannot be conceived as stratified, and that the term "phenomenon" has to be meant as "illusory appearance" in a supposedly Kantian sense. At the same time, I will discuss Husserl's conviction that the 'technization' of science entails a philosophical loss of meaning of the scientific image of the world.

Keywords: life-world, quantum physics, scientific realism.

The aim of this article is to assess whether the notion of "life-world" could be helpful for a philosophical theory that assigns a primacy to the scientific view of the world when it comes to establish what exists, and, accordingly, what "reality" means. In establishing this parallel, I am referring, on the one hand, to one of the most important concepts of Edmund Husserl's late phenomenology (the concept of "life-world"); on the other, I am trying to endorse the point of view defended by Wilfrid Sellars in his famous study *Philosophy and the Scientific Image of Man in the World*.

My attempt is to integrate two views that scholars have routinely described as conflicting.[2] This does not mean that the perspectives of the two thinkers

[1] I am very thankful to Daniele De Santis, Agnese Di Riccio, Roberto Gronda, and Michele Mancarella for their precious comments on a previous version of this article; my special thanks to Giacomo Turbanti, for his comments and for the intellectual exchange that gave rise to this volume and which, I hope, will give rise to new projects.

[2] This assumption is shared by Christias (2020; 2018), who argues that the categorial framework of the life-world entails a scientific instrumentalism which takes the unobservable objects and properties of science as mere "calculational devices," and Sachs (2020; 2014), who holds that Husserl falls

ISSN (print) 2281-8618, ETS doi: 10.4454/philinq.v10i1.387
Submitted: March 2021
Accepted: October 2021

I have just mentioned will be accepted in full in this article. However, I think that the integration I am going to argue for is, if not the only possibility, at least one of the most effective ways to explain how a philosophical perspective that advocates for a peculiar version of scientific realism can take into account the conceptual framework of persons.

I will articulate my argumentation into three parts. In the first, I will show that a theory of the life-world does not necessarily entails the endorsement of the primacy of the manifest over the scientific image of the human beings in the world. In the second, I will consider the image of the world that derives from the quantum revolution in physics in order to challenge two significant assumptions made by Sellars: the first is that reality cannot be conceived as stratified, the second is that the term "phenomenon" has to be understood exclusively in the supposedly Kantian sense of "illusory appearance". In the third section, I will start by challenging Husserl's conviction that the 'technization' of science entails a philosophical regression of the scientific image of the world; this will allow me to take into account the metaphilosophical implications of a perspective that assigns a role to a theory of the life-world for the justification of a form of scientific realism. Indeed, this study adopts a metaphilosophical approach for at least two reasons: firstly, it is an investigation that reflects on the meaning and conditions of a specific philosophical perspective, such as the theory of the life-world; secondly, this reflection aims to express itself on the task of philosophy in general.

1. *The conflicts of images*

As is widely known, in *Philosophy and the Scientific Image of Man* Sellars maintains that, in search for a unity of knowledge, philosophers are necessarily confronted "not by one complex many-dimensional picture," but "by two pictures of essentially the same order of complexity, each of which purports to be a complete picture of man-in-the-world" (Sellars 1962: 4). Initially, Sellars designates these two conflicting views as the manifest image and the scientific image. He identifies the former with a refinement, or a sophistication, of the original way in which humanity refers to objects in the ordinary life, whereas the latter view is that which has been developed from the modern scientific revolution.

At a second stage, Sellars acknowledges that the manifest image is scientific in a peculiar sense: it proceeds indeed from the standard ways in which

under the Myth of the Given. De Santis (2020) and, earlier, Soffer (2003) exclude this risk. Finally, let me also refer to my article (Manca 2021), where I argue that the point of divergence between Sellars and Husserl does not lie in the conception of the Given but in their different way of understanding the spontaneity of thinking.

objects appears in the perceivable world, and, by induction, infers the universal forms of the phenomenal realm. By contrast, the framework that he had initially defined as "scientific" might more accurately be called "the postulational or theoretical" (Sellars 1962: 7). It is determined by the only procedure that the inductive approach of the manifest image cannot take into account: it postulates "imperceptible objects and events for the purpose of explaining correlations among perceptibles" (Sellars 1962: 19).

Both conceptual frameworks pretend to deal with the only effectively real world. This is the core of their clash. In order to describe their differences and take a position, Sellars distinguishes three possible lines of argument:

> (1) Manifest objects are identical with systems of imperceptible particles in that simple sense in which a forest is identical with a number of trees.
> (2) Manifest objects are what really exist; systems of imperceptible particles being 'abstract' or 'symbolic' ways of representing them.
> (3) Manifest objects are 'appearances' to human minds of a reality which is constituted by systems of imperceptible particles. (Sellars 1962: 26)

Those who advocate the primacy of the manifest over the postulational image of the world could opt for the first two lines of argument.

As regards thesis (1), Sellars argues that there is nothing immediately paradoxical in thinking that "an object could be both a perceptible object with perceptible qualities *and* a system of imperceptible objects, none of which has perceptible qualities" (Sellars 1962: 26). This is why systems can effectively display properties that their parts do not have. A condition for defending this position is to recognize that the so-called emergent properties of a system depend on the properties of, and relations between, its constituents. Once we accept this, we are directly brought to endorse thesis (Sellars 1962: 3). Indeed, "if a physical object is *in strict sense* a system of imperceptible particles, then it cannot as a whole have the perceptible qualities characteristic of physical objects in the manifest image" (Sellars 1962: 27). Thus we must conclude that "manifest physical objects are 'appearances' to *human perceivers* of systems of imperceptible particles" (Sellars 1962: 27).

On the contrary, thesis (2) inevitably conflicts with the scientific realist's view, insofar as it identifies the objects of experimental natural sciences with instruments, or constructs, for explaining specific aspects of reality but these instruments can in no way grasp the essential features of the world as it is in its concreteness and wholeness. Scholars who, from a Sellarsian standpoint, have discussed the perspective Husserl elaborates in his latest masterpiece, *Crisis of the European Sciences*, assume that Husserl would have endorsed this second

line of thought.[3] We cannot exclude that there are good reasons for doing this, but this view is more controversial than one might initially think.

In the reconstruction of Galilei's foundation of modern physics – where Galilei stands more for an entire movement rather than for a historically determined individual –, Husserl demonstrates how the difficulty of mathematizing sensible qualities (colours, tones, and so on) brought the new physics to express discovered laws through algebraic formulas. In his view, this resulted into an automatic procedure that progressively concealed the original picture of the world from which the revolution moved (nourished by the Platonic impulse to seek the ideal forms of the world): "In algebraic calculation, […] one calculates, remembering only at the end that the numbers signify magnitudes. Of course one does not calculate 'mechanically', as in ordinary numerical calculation; one thinks, one invents, one may make great discoveries – but they have acquired, unnoticed, a displaced, 'symbolic' meaning" (Husserl 1954: Eng tr. 44-45).

Here Husserl does not assert that systems of imperceptible particles are 'abstract', or 'symbolic', ways of representing manifest objects. He is rather claiming that algebraic formulas are symbolic constructs for grasping the laws of natural processes. Indeed, Husserl does not deny that what we perceive as colour is the result of stimulation of photoreceptor cells by electromagnetic radiation. He limits himself to point out that this was not taken for granted at the time of Galilei – evidently with the sole exception of those who fully advocated the alternative (see Husserl 1954: Eng. tr. 36-37).

Therefore, Husserl is not opposing the picture of the world arising out of the modern revolution. More simply, he is interested in highlighting the decisive support that algebra gave to the scientific enterprise, while claiming that in the history of classical physics the use of algebra favoured a 'technization' of the method of investigation. In other words, the revolutionary discoveries of early modern thinkers such as Galilei turned what was previously only a method (i.e. an art of measuring) into the proper object of investigation (i.e. in what has to be known).

More explicitly, Husserl does not think that imperceptible particles do not ultimately exist. If that were the case, Husserl's view would be as ingenuous as that of the philosopher who believes that manifest objects do not exist. Husserl holds that systems of imperceptible particles can legitimately claim a specific degree of reality, just as manifest objects are entitled to be recognized as an-

[3] As is widely known, Sellars was indirectly influenced by Husserl's phenomenology via Farber. In his 1962 essay, Sellars does not explicitly mention Husserl's life-world, but he does that in a later essay, in which he argues that "the manifest world—the *Lebenswelt*—has its own intelligibility," but it also "poses questions which it does not have the resources to answer" (Sellars 1981a: 282).

other degree of reality.[4] This lead us to reconsider from Husserl's perspective the first of the lines of thought Sellars outlines.

I am quite sure that a reader of Husserl would be struck by the condition that Sellars introduces when he reduces the first to the third line of argument. Indeed, he clarifies that the condition for accepting that a system can have properties that its parts do not share would be taking as "paradigm example" for that "the fact that a system of pieces of wood can be a ladder, although none of its parts is a ladder" (Sellars 1962: 26). According to the phenomenological perspective, an organism cannot be reduced to its material constituents (contrary to an aggregate such as a ladder) without losing some of the properties that pertains to it as a whole. This would be the case of the forest, to which Sellars refers in order to formulate this first line of thought. Although this aspect marks a divergence between Husserl and Sellars, it does not seem to me to be the fundamental breaking point between them. Indeed, Sellars grants that the properties of a system do not simply consist of the properties of its constituents, but also the properties that arise out of the relations between the constituents. For instance, the ecological resilience of a forest is clearly determined by the way in which its trees interact, insofar as they live as a whole.

More radically, the original element of divergence between the two perspectives lies in the fact that Sellars considers the manifest image of the world inadequate and unacceptable as regards the "account of what there is *all things considered*," and this in spite of its adequacy "for the everyday purpose of life" (Sellars 1962: 27). To his view, "the world of everyday experience is a phenomenal world in the Kantian sense" (Sellars 1968: 173), for manifest object are just illusory appearances that conceal the real world of imperceptible entities. Differently from Husserl, Sellars not only excludes from the onset the possibility of accepting a stratified view of the real world as constituted by different degrees of reality, each governed by its essential laws; he also denies that the notion of "phenomenon" could have different meanings. Thus, Sellars's scientific realism rules out the possibility of a philosophical account of the life-world, unless one describes it as a good way to accounting for how persons see the things from within of the illusory framework in which they straightforwardly live.

However, this position comes at a price, and Sellars seems to be completely aware of this.

[4] See Kerszberg 2012 for an overview of Husserl's conception of science in light of his theory of the life-world, and Trizio 2021 for an accurate investigation into the relations between phenomenology and natural sciences.

In the middle of his essay, exactly before delving deeper into the features of the postulational image of the world, Sellars identifies his "primary concern" with the following question: "In what sense, and to what extent, does the manifest image of man-in-the-world survive the attempt to unite this image in one field of intellectual vision with man as conceived in terms of the postulated objects of scientific theory?" (Sellars 1962: 18). He suddenly acknowledges that "to the extent that the manifest does not survive in the synoptic view, to that extent man himself would not survive" (Sellars 1962: 18). In other words, if there was no room for a theory of the life-world, the conceptual framework of persons could not be adequately investigated and understood.

At the end of the essay, the exclusion of both a stratified theory of what really exists and the absence of a multi-faceted notion of "phenomenon" leads Sellars to conclude that the conceptual framework of persons "is not something that needs not to be *reconciled with* the scientific image, but rather something to be *joined* to it" (Sellars 1962: 40).

In light of this, the issue I would like to tackle in the following two sections is whether we really have to be content with the perspective according to which the clash between the manifest and the scientific images of man in the world can be transcended only in imagination, as Sellars suggests. In order to pursue this goal, I would like to bring into play the picture of the world and of the scientific enterprise fostered by the quantum revolution.

2. *A complicated tissue of events*

As is widely known, the quantum interpretation of microscopic processes in physics was elaborated in the first half of 20th century through the collective effort of different personalities following different (and sometimes opposed) trajectories. Over the last decades, this arduous theory has known a great experimental success, but the ontological implications of its assumptions are the subject of a wide-ranging and multifaceted debate. In this section, my intention is neither to discuss this debate nor to provide a personal interpretation of the ontological meaning of quantum theory. Rather, I will try to show that, according to the image of science that the standard "Copenhagen" version of quantum revolution provides, the impression that Sellars sided with the worldview defended by physics insisting on the clash between the manifest and the postulational images, while Husserl strengthened the position of common sense through his theory of the life-world is questionable to say the least. In the first part of the section, I will go over the key, initial, stages of quantum revolution; I will then move to a discussion of some

of its most significative achievements and their relevance for the integration I am proposing.[5]

Quantum physics developed in an attempt to explain the wave-particle duality, i.e. the fact that every microscopic event can be investigated now by describing the objects involved as particles, now by comparing their behaviour to that of a wave.

In 1900, Max Planck was seeking to understand why experimental results show that energy in a black body is distributed over various wavelength ranges when he boldly guessed that the energy carried by an electromagnetic wave comes in lumps. In the article that would win him the Nobel Prize in 1921, Einstein (1905) applied Planck's lumpy picture of wave energy into a new description of light, introducing the concept of a stream of tiny particles then called photons in order to explain the strange features of the photoelectric effect (i.e. the fact that when electromagnetic radiation shines on certain metals they emit electrons; surprisingly, this does not depend on an increase in light intensity; it is the colour of the impinging light that determines whether or not the electrons are ejected, and if they are, the amount of energy they have). Planck's and Einstein's insights marked a great progress in the history of physics but they also sparked a heated debate. Indeed, their results suggested that light has particle properties – a view supported by Millikan's experiment and the discovery of Compton scattering –, while other experiments demonstrated that photons somehow embody wave-like features of light (this is the case of Young's experiment, now known in a particular variant as the double-slit experiment).

Niels Bohr appropriated Planck's insight in order to elaborate a model for an atom. He imagined that each electron does not orbit the nucleus freely but as held in place; it vibrates back and forth by carrying energy only in multiples of some basic "quantum," that is, by taking on only a limited set of values. Arnold Sommerfeld, from the University of Münich, further enhanced this depiction by demonstrating that the ellipticity of electron orbits was quantized as well.

On the one hand, the Bohr-Sommerfeld model for an atom excited many young scholars, like Pauli and Heisenberg, who decided to explore the issue further. On the other, it raised a certain concern, voiced by figures such as Einstein and Ernest Rutherford, who in 1911 had found out the nucleus of the

[5] For my reconstruction, I followed Greene 1999, Healey 2017, Lewis 2016, and Lindley 2007. See also Maudlin 1999 and van Fraassen 1980 for an interpretation of quantum mechanics in comparison with Sellars's distinction between two images of the world. Finally, the encounter between phenomenology and quantum physics has a long history; for an up-to-date, multifaceted reflection on the points of contact between the two ways of depicting the world see Bitbol 2020, de La Tremblaye 2020, French 2002 and 2020.

atom by working on radioactive alpha emanations. Rutherford asked in a letter to Bohr (dated March 20, 1913) how an electron decides with what frequency it is going to vibrate and when it passes from one stationary state to another (see Bohr 1981: 583). Einstein complained that Bohr's theory of a quantum jump that brings an electron to an abrupt transition from one energy level (or orbit) to another strongly questioned the importance of causality in physics. This jump happens indeed without any identifiable cause, but with a spontaneity that resembles that of the radioactive decay of a nucleus. This is why, in a letter (dated April 29, 1924) to Max Born from the University of Göttingen and his wife, Einstein wrote that he might as well have been "a cobbler, or even an employee in a gaming-house" rather than a physicist (see Einstein, Born M. and H. 1971: 82).

And yet, the young French scholar Louis de Broglie applied Einstein's depiction of light as a stream of particles to the Bohr-Sommerfeld atom. He calculated that a wavelength is exactly equal to the orbit's circumference. This allowed him to put forward the hypothesis that wave-particle duality – the feature that Einstein ascribed to light – is a characteristic of matter as well (see de Broglie 1925). Around the same time, Werner Heisenberg (who had studied with Sommerfeld in Münich, Born in Göttingen, and finally Bohr and his collaborator Kramers in Copenhagen) was working on how electrons in atoms behave. To explain the discontinuity of the inner activity of the atom, Heisenberg connected Kramer's hypothesis and Born's proposition. Kramer maintained that the structure of the atom was closer to a set of tuned oscillators rather than the solar system (in which electrons follow well-defined orbits, governed by classical mechanics). Born, on the other hand, suggested to substitute classic differential calculus (elaborated independently by Newton and Leibniz to deal with continuous variation and incremental changes) with a mathematical system that had the difference between the stages at its basic elements, rather than the states themselves.

With Pascual Jordan and Born, Heisenberg formulated the fundamental equations of the new mechanics by employing the so-called matrix method, based on multiplying two list of numbers together (see Heisenberg 1925; Born, Jordan 1925; Born, Heisenberg, Jordan 1925). In the meantime, inspired by De Broglie's calculus, Erwin Schrödinger (1926a) achieved similar results by working on an equation for describing how a matter wave should evolve. The particle-wave duality seemed to be transposed at the level of mathematical calculation. However, Schrödinger himself demonstrated the equivalence of the two calculation systems (see Schrödinger 1926b), while Born (1926) proposed to interpret an electron wave in probabilistic terms with the intention of avoiding a mathematical duality: the wave is largest where the electron is most likely

to be found, and progressively smaller in locations where it is less likely to be found. Paul Dirac (1928) definitively incorporated matrix mechanics and the Schrödinger equation into a single formulation.

The height of quantum revolution was reached with Heisenberg's formulation of the so-called uncertainty principle in a 1927 article, entitled *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*. Here Heisenberg aims to define some terms of classical physics which remain valid in quantum mechanics and, by so doing, showing that the relative quantities "can be determined simultaneously only with a characteristic indeterminacy" (Heisenberg 1927: Eng. tr. 62). He came to this conclusion by introducing a criterion of observability. In his previous article *Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen*, he had already postulated that a basis for theoretical quantum kinematics and mechanics could be founded "exclusively upon relationship between quantities which in principle are observable" (Heisenberg 1925: Eng. tr. 261). In his 1927 article, Heisenberg reverses the perspective of classical physics—the foundational role that a particle's position and velocity played for mechanics—, moving from the observation by measurement of the frequencies and intensities of the fluctuations of the atom, and inferring the position and velocity of electrons only on the basis of those. The point is that the experiments that can help us to measure, for instance, the position of the electron inevitably interfere with the event they are determining, making it impossible to measure the velocity with a similar accuracy.

For example, when we illuminate the electron under a microscope, every observation of scattered light coming from the electron presupposes the Compton effect:

> At the instant when position is determined – therefore, at the moment when the photon is scattered by the electron – the electron undergoes a discontinuous change in momentum. This position is greater the smaller the wavelength of the light employed – that is, the more exact the determination of the position. At the instant at which the position of the electron is known, its momentum therefore can be known up to magnitudes which correspond to that discontinuous change. Thus, the more precisely the position is determined, the less precisely the momentum is known, and conversely. (Heisenberg 1927: Eng. tr. 64)

This should be enough to understand the revolutionary achievement represented by the quantum interpretation of microscopic physical events. Every experiment for the determination of a feature of the electron inevitably perturbs or alters the atom itself. Indeed, we can predict with accuracy the probability that an electron occupies a particular position, but at the cost of admitting that it is impossible to determine velocity with the same accuracy (and vice

versa), or to determine where the electron is, or even how the elements of the atom behave before getting to measurement.

I think we can end our reconstruction at this point and turn to the theoretical consequences that are relevant to our goals. I see four, strictly interconnected, aspects which needs to be assessed. Let me summarize them as follows:

1) A weakening of strong realism in science, in light of the argument that measurement interferes with the event it aims to determine;
2) A rejection of the materialist ontology that characterizes modern classical physics;
3) A decisive enhancement of the constructive capacity of mathematics;
4) An extension of the concept of intuition.

The first two aspects affect Sellars's description of the scientific image, while the last two question some of Husserl's assumptions. In the last part of this section I will focus on the first two; in the next section, I will try to clarify how the last two points leave space for a theory of the life-world within a view that ascribes to science a decisive role in the definition of what exists.

In light of the quantum scientific image of the universe, it is difficult to follow Sellars's focus on what exists. More accurately, we should speak of what happens when we endeavour to observe it. Heisenberg reminds that "in the drama of existence we are ourselves both players and spectators" (Heisenberg 2000: 25). We cannot escape from this ambiguous condition. Thus what we observe in physics – thanks to the measuring device that we constructed – is not "nature in itself but nature exposed to our method of questioning" (Heisenberg 2000: 25).

This formulation seems to suggest a form of anti-realism in science, but we should more appropriately speak of a pragmatic realism:[6] indeed, it would be misleading to think that there is no quantum world; the physics does describe the nature of the microscopic universe. The point is rather that we can only grasp it through an abstract description. While the references are real, the correlated event of our description happens at the microscopic level in a way that is at odd with our ordinary way of seeing the world. Conversely, the tools we use for elaborating our description are inevitably mediated by our mental constructs, by our capacity of devising signs that can designate the event as accurately as possible, or by figuring out appropriate experiments and constructing effective devices to verify our hypotheses.

This leads us to the second point. Heisenberg maintains that "in classical physics science started from the belief – or should one say from the illusion? – that we could describe the world or at least part of the world without any

[6] On Heisenberg's pragmatic realism see the interesting reflections of Cappelletti 2001.

reference to ourselves," but this belief only results from "a refinement of the concepts of daily life" (Heisenberg 2000: 23). We can easily separate the object from the subject only by confining ourselves to the macroscopic sphere, in which we straightforwardly live. In the microscopic one, however, the interaction between the system and the observer is inescapable. In 1996, Carlo Rovelli proposed a refinement of the standard "Copenhagen" depiction of the state of affairs. In the traditional view of quantum physics, the observer who makes a measurement on a quantum system is macroscopic. On the contrary, in the relational interpretation of quantum mechanics, by using the word "observer" we should not refer to a "conscious, animate, or computing […] system;" rather, in a more Galilean fashion, Rovelli identifies the observer with "any physical object having a definite state of motion" (Rovelli 1996: 1641). As one can easily get, in this way the notion of "observation" merges into that of interaction: "any system, irrespectively of its size, complexity or else, can play the role of the quantum mechanical observer" (Laudisa, Rovelli 2019). The point is no longer what can I grasp with the help of technology and following my theoretical hypothesis, but the extent to which a system is able to affect the other.[7]

Quite surprisingly, when seen from the quantum perspective, Sellars's scientific image turns out to be a refinement of the manifest image, which is still more sophisticated than Husserl's ontology of the life-world. Sellars thinks in terms of a world that exist below and behind the manifest world. Even though we can grant him that his depiction of the real world as a system of particles is only an approximative way of describing a more complex universe – characterized by fields, forces, a matter that behaves like a wave – Sellars's assumption that the representation of a world in terms of a system of imperceptible objects is completely independent of the subject that investigates it is absolutely questionable.

Starting from a position that is the opposite of Husserl's, quantum theory paradoxically encroaches the same two objections that we raised against Sellars by following Husserl's perspective at the end of the previous section. More

[7] In order to fully address the question of whether realism plays a conceptual role in quantum physics, and so understanding how the theory changes on the basis of different interpretations of the notion of "observer" itself, one would need to extend the investigation to what has been referred to as the "second quantum revolution" (see Maudlin 2019 for a depiction of the latest outcomes of quantum revolution, and Maudlin 1999 for an interpretation of the second quantum revolution in light of Sellars's distinction of two images). This move was made necessary by Einstein, Podolsky and Rosen's attempt to demonstrate the incompleteness of standard quantum theory (see Einstein, Podolsky, Rosen 1935), and by Bohm's introduction of the topic of hidden variables (see Bohm 1952). It was in fact Bell's demonstration to suggest that quantum theory can be considered complete if it accepts the non-locality of quantum systems, i.e. the possibility of systems interacting at a distance (see Bell 2004, and Maudlin 1994).

explicitly, the two points that Sellars misses are (1) that a scientific image of the world must elaborate a stratified conception of what is real, and (2) that "phenomenon" does not only coincide with "illusory appearance".

In a supplementary text to *Crisis* Husserl briefly dwells on the meaning of quantum physics for the history of scientific knowledge. Here he appreciates the probabilistic and relativistic approach it adopts, and insists on the fact that the totality of the world is divided into fields with different typical structures (see Husserl 1954: 389-390).[8]

In other words, in physics we have to think in terms of scales. When Planck supposed that an electromagnetic wave comes in lumps, he found out that the minimal unit of energy a wave can carry is proportional to its frequency. This led him to introduce a new constant in physics for expressing this quantum phenomenon. It is now known as Planck's constant, denoted "h" and, in the Dirac's reduced formulation, "ħ" (pronounced "h-bar"). It is extremely small – it is about a billionth of a billionth of a billionth in everyday unit (more specifically it is 1.05 x 10-27 grams-centimeters2/second). This entails that in those layers of reality in which Planck's constant is irrelevant, quantum effects are not significative. This is the case of the manifest world.

Richard Feynman, the famous American physicist who worked on a theory of quantum electrodynamics and invented the mathematical path integral formulation of quantum mechanics, reinterpreted the double-slit experiment by arguing that the probability for an electron to arrive at a point on the screen is built up from the combined effect of every possible way of getting there. Accordingly, he suggested that particles must be viewed as travelling from a location to another along a set of infinite trajectories. However, he also showed that if we examined the motion of macroscopic objects with his new method of calculus, all paths but one cancel each other out when they are combined. The trajectory that remains valid is approximately the one emerging from Newton's law of motion. This explains why in the everyday experience a manifest object follows a unique, predictable trajectory when travelling from a location to another.[9]

This is further evidence for questioning Sellars's view of a world behind the other, and the following identification of the manifest world with the phe-

[8] As Argentieri (2009) has argued in his illuminating essay, Husserl's conception of a stratified reality is very close to the ontological perspective Heisenberg defends in an essay published posthumously, in which reality, understood as the totality of connections of life, is ordered "in diverse areas" (Heisenberg 1989).

[9] See Feynman 2010, ch. 19: 4: "As we apply quantum mechanics to larger and larger things, the laws about the behavior of many atoms together do *not* reproduce themselves, but *new laws*, which are Newton's laws, which then continue to reproduce themselves from, say, micro-microgram size, which still is billions and billions of atoms, on up to the size of the earth, and above". Cf. also Feynman 2006 and 1983.

nomenal realm in a Kantian sense. Moreover, if the world of imperceptible microscopic objects were the only real one, could be said about the nature of astronomical phenomena? Does Einstein's revolutionary depiction of four-dimensional spacetime refer to something that can be taken to be real? In this case as well, since the manifest objects in our perceivable world are very far from moving near light speed, the relativistic effects described by Einstein are so irrelevant to be undetected.[10]

To conclude, in contrast with Sellars's (approximative, in light of our argument) identification of reality with a system of imperceptible entities, we might quote Heisenberg's appropriate definition. By noticing that the world cannot be divided into "different group of objects but into different group of connections," Heisenberg suggested that "the world thus appears as a complicated tissue of events, in which connections of different kinds alternate or overlap or combine and thereby determine the texture of the whole" (Heisenberg 2000: 64).[11]

3. *The space of persons*

Husserl's belief that the algebraization of science entails a gradual loss of meaning generated among some interpreters the erroneous conviction that in Husserl's view the perceivable world is the only source of meaning. On the contrary, since the *Logical Investigations*, following Bolzano and Lotze, Husserl insists on the objectivity of meaning, i.e. on the fact that it is valid independently of its occurring in a psychic lived experience or being verified in a sensible experience.

It is true that Husserl sometimes exposed his notion of a "life-world" to the risk of being identified with a closed set of objects and behaviours closely linked to perceptual experience. Yet I think that this identification of the life-world with the perceptual realm is merely a remnant of the original definition of the notion Husserl offered at the end of the 1910s, but which no longer holds for the context of *Crisis* we are here considering.

At first, Husserl used the term "life-world" to designate the sphere of pre-predicative experience as opposed to that of judgement. However, his focus shifted over the years. As Iso Kern has clearly pointed out, "if it was initially formulated as a problem concerning foundational relationship between the scientific concept and the preconceptual intuition, in the course of his reflec-

[10] Reasoning in terms of scale in physics is crucial to avoid the conflict that would otherwise arise between quantum mechanics and theory of relativity (see Greene 1999, ch. 5).

[11] After all, Sellars seems to be drawn in this direction in his famous Carus lectures (Sellars 1981b), in which he argues that a positive ontology is only possible if we speak in terms of pure processes.

tion it was transformed into the problem concerning the fundamental relation between the abstract world of objective theory and the concrete, historical world of subjective life in which 'theoretical praxis' belongs as one mode of human praxis *among others*" (Bernet, Marbarch, Kern 1989: Eng. tr. 222).[12]

When examining the process of mathematization of nature that started new physics (now labelled as classical physics), Husserl recognizes that "the process whereby material mathematics is put into formal-logical form […] is perfectly *legitimate*, indeed necessary," but he also adds that "this can and must be a method which is understood and practiced in a fully conscious way" (Husserl 1954: Eng. tr. 47). In other words, in Husserl's view the tendency to employ *formulae* is perfectly justified by the fact that they facilitate our capacity for predicting what is to be expected in experience or in the experimental verification. However, as we have already emphasized, Husserl sees the process of algebraization of science as an authomatization of thinking that leads to a systematic displacement of the symbolic method of calculation, which progressively becomes the very object of investigation instead of maintaining its necessary but only supportive role for knowledge.

This shows that what is at stake for Husserl here is a historical process. The loss of meaning technization would imply does not merely concern the way in which the 'world of formulae' is rooted in the perceivable world. It rather concern the capacity of the mind to lead the symbolic transformation of scientific thinking back to a historical change of paradigm, thus back to a cultural event determined by an epoch's worldview, and located in what Sellars would call the space of reasons—the conceptual framework of persons where we demands for and articulate our reasons.

In *Crisis* and in the famous, posthumously published, essay on the origin of geometry, Husserl stresses that the life-world is historical, thus it has to do not only with perceptible objects but also with the culture we inherit, with the values we embrace or reject, with the language of communities and individual intentions (to state this in Sellarsian terms again). The natural attitude we straightforwardly adopt within the horizon of the life-world coincides with the personalistic attitude that characterizes us as citizens who hold a certain worldview and are guided by cultural motivations. This explains why some of the discoveries of the modern scientific revolution have become so familiar to us: even if we do not study them in depth, we would never dream of questioning their soundness. For instance, even though the earth appears to us perceptually as the stable ground of our everyday experience, we have

[12] For a thorough reflection on the development of Husserl's conception of the life-world let me refer to Manca 2016, section 1, ch. 3.

become familiar with the awareness that it revolves around the sun.[13] This implies that objects of a life-worldly view of things can be not only what one grasps or verify by perceptual experience, but also knowledge that we acquire, theories that we postulate, discoveries that become the common background of an epoch.

In this sense, when Husserl identifies the loss of meaning brought about by the technization of science with the concealment of the life-world understood as the ground of each cultural event (like the introduction of a new method into natural science), he is pointing out to the difficulty for the community to recognize a decisive advancement in knowledge it nonetheless accomplished.

When Sellars describes the manifest image as a refinement of the original one, he emphasized that in such a scientific elaboration of the ordinary depiction of the world the categorial conditions do not effectively change, but all object are assimilated to persons, albeit some of them as "truncated". Husserl's scepticism about the process of technization in the scientific realm turns the attention to the acts more than on the objects. According to this noetic perspective, the change of paradigm was effectively generated by a truncation of the social, cultural, and more generally shared experiences, rooted in the space of persons.

The problem seems to coincide with the fact that natural scientists focused on the elaboration of calculus as a method rather than delving deeper into the philosophical reasons behind it. For sure, this has to do with the fact that philosophy has progressively lost the primacy in the elaboration of an image of the world that it held for centuries, a primacy that was assigned to natural sciences.

It is misleading to think that the specialisation of knowledge leaves little room to the reflection on the philosophical implications of scientific discoveries. Indeed, in the 20th century thinkers such as Einstein, Bohr, and Heisenberg not only gave conspicuous space to the need of elaborating a conceptual framework of the world, but in some cases they started from a philosophical

[13] It is not by chance that I cite this example. There is a manuscript by Husserl, further valued by Merleau-Ponty, in which Husserl argues that even if, after Copernicus, the earth must be considered one body among others, it remains for us the firm ground of our experience (see Husserl 1968). With this, Husserl does not want to return to the Ptolemaic geocentric theory. Rather, he wants to argue that the "paradigm shift" in the context of the scientific image does not affect the manifest image. The "Earth" as the object of a theoretical-objective intention is completely different from the earth as the reference context of ordinary sensory experience. This explains why we are able to become familiar with the cognition that the earth revolves around the sun and that the apparent movement of the sun is only an illusion determined by the movement of the earth around its own axis, while perceiving the earth as the stable ground of our experience. Let me refer to my essay (Manca 2014) for an account of Merleau-Ponty's reading of this manuscript, and his picture of science in comparison with Sellars's.

reflection on the existing paradigm to formulate their hypotheses – that is, to postulate an image of the universe whose verification depended on experiments frequently not conducted by them.

Yet, I think we should look at another, more decisive point: these thinkers carried out their theoretical activity in and through a complex mathematical language. Even though later in their research activity they devoted themselves to disseminating their discoveries in ordinary language – obviously leaving behind some of the rigour of the proper formulation–, mathematics remains the language they developed not simply to communicate the conceptual contents of their theoretical activity, but also to conceive them.

For Einstein (1931: 69), Newton's choice of enunciating the laws of motion in the form of total differential equations was "perhaps the greatest intellectual stride that it has ever been granted to any man to make". Newton makes only a marginal, not systematic use, of symbolic, mathematical language, but this paradoxically allowed his successors to elaborate a scientific paradigm in opposition to that which Newton himself consolidated.

Thus, mathematical language leads human thinking to refine its art of measuring, independently of the capacity of speaking of what is real. For instance, some of the most influential representatives and interpreters of quantum mechanics look at wave function – which describes the quantum state of an isolated system – exclusively as a calculational tool that refers to an abstract mathematical space, but cannot be said to have an objective, physical existence.

Husserl's suspicions aside, these two examples proves that he is right in identifying mathematics with a human construct for grasping the essential features of the physical world. In other words, from this perspective, it is not possible to maintain that nature is written in hidden mathematical characters, which human mind would then need to decipher (as Galilei had it), because the category of subject cannot be separated from that of object. We have seen that this thesis is common to quantum theory and Husserl's phenomenology. Mathematics, on the other hand, is human mind's most effective tool for forcing nature to speak; it is the linguistic practice that allows human mind to delve deeper into its interaction with nature, with which it actually shares an inborn unity. This seems to me the only way to hold together Husserl's conviction that mathematics alone can depict the "true being-in-itself" of nature and his critique of the unreflective use of algebraic formulas, which natural scientists pursue in even more systematic ways.

Moreover, this conception of mathematics reveals the point of contact (and, in fact, also the point of divergence) of Husserl's idealism and Sellars's nominalist psychologism. As is widely known, one of the most famous theses that Sellars (1956: 162) defended in his *Empiricism and the Philosophy of Mind*

is "the denial that there is any awareness of logical space prior to, or independent of, the acquisition of a language". Even though Husserl insists that rational rules are already in play at the level of passive, unconscious experiences of the world, and that sense displays a pre-expressive dimension, he would agree with Sellars's belief that language is the natural form of sense. Thus, for both thinkers, the awareness of concepts drawn on in our life is strictly dependent on the use of a language.[14] In the case we are examining, the mind could not acquire awareness of the quantum image of world without the use of the mathematical language. Neither the relativity of time and space, nor the equivalence between matter and energy, nor the fluctuations of the atom, nor the ambiguous behaviour of matter in general, would be accessible to us without relying on an advanced mathematics.

As Husserl explains in *Crisis*, at the basis of modern scientific revolution there is the conviction that mathematics is no longer a formal transposition of concrete relations, but a new way of thinking. Algebraic formulas are not simply abbreviations; rather, they stand for an ideal construction of the world. The technization of science shows that the conception of mathematics as an ideal garment of reality, as something that is built on the life-world, cannot work. The step further that the quantum worldview takes is to understand mathematics as a linguistic practice, which allows not simply to describe but to organize a scientific experience and to construct (via technological mediation) a connection – beyond the ordinary one – between that event which is called "mind" and that event which is the phenomenon "world".

We can thus move to the second aspect I wanted to highlight. Heisenberg titles his 1927 article on uncertainty in quantum physics *Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik*. Whereas the adjective "anschaulich" has been translated as "physical," and more seldomly as "actual," it literally means "intuitive". Heisenberg is here clearly referring to what can be measured and not to what can be perceived by the naked eye. The terms "physical" or "actual" are misleading because they could suggest a reference to a content that has to be taken as real independently on the observer who grasps it ('observer' should be understood according to the extensive meaning we have already expanded upon by following Rovelli). By contrast, as is already pointed out, quantum theory takes the product of the interaction between physical systems as always known. Thus, if one looks for a more explanatory translation, closer to Heisenberg's own intention, the adjective "anschaulich" could be rendered as "graspable". Indeed, Heisenberg would prefer

[14] Instead, the difference between the two thinkers is that, for Sellars, a concept is a social, intersubjective construct of ordinary language, while, for Husserl, a concept always refers to an ideal sense.

to emphasize the need to start from what can be observed in an experiment in order to determine what the theory can grasp or not (and to what extent), rather than making use of traditional concepts without discussing them.

Still most significantly, the use of the term "anschaulich" extends the notion of intuition that is not that far to the one that Galilei carries on with his revolution in physics. Like Heisenberg did with respect to the tradition of the classical physics that stems from Galilei – and in contrast with the Aristotelian conceptual apparatus that he inherited from the Scholasticism of his time – Galilei himself discloses the intuitive power that a theoretical attitude offers in relation to reality. Imagining a mental experiment, formulating hypotheses, constructing an experimental verification of them, grasping the ideal forms of a concrete phenomenon, and using mathematics for this, are closely interconnected activities.

All this shows us that the concept of "intuition," as it is the case for that of "meaning," to speak more generally, develops in history. This does not entail that the meaning of a concept is contingent and can be modified at will. Rather, this suggests that one's life-world can always be extended to new categories, to new way of grasping a reference for what we mean.

Apart from the cases in which Husserl risks to completely identify the life-world with the perceivable dimension of reality, he frequently describes the life-world as the open horizon within which we can familiarize with new objects, while others may be reconsidered to the point of losing their original grip; as the context in which new attitudes may be undertaken, old practices forgotten, or, in Husserl's words, allowed to be sedimented.

This brings us to the conclusive remarks of this essay. A philosophical consideration of the life-world seems to be necessary even for a perspective that ascribes a primacy to science in outlining what is real. Indeed, Husserl's life-world coincides with the sphere of persons, in which we articulate and are able to justify our reasons, i.e. to prefer an image of the world to another. The involvement of quantum theory led us to a further reflection on this issue. When we state that we should replace the belief that reality is a system of imperceptible objects with one that depicts it as a complex tissue of events and interactions between systems (on the basis of a scientific perspective that has enjoyed enormous experimental success), we are not modifying the belief that the best way to understand what is real is to postulate what can be observed by measurement. In putting forth this option, however, we are removing an obstacle in understanding the conflict between the two images. We are, in fact, understanding that the conflict is rooted in a historical paradigm that develops a certain representation of the terms at stake. In other words, this position emphasizes that the point of contact between the two images does not lie in

making perceivable what is indeed not perceivable, but in the ability of getting acquainted with a vision of the world that could only be partially expressed in ordinary language. An alternative is possible that avoids the conflict ending in an aporia.

The option we are offering does not only allow to provide a naturalistic, scientific, description of the form of life we identify in the manifest image with the term 'person' so that a conjunction of the two perspectives is made possible through an analogy that links one category of the scientific image to another in the manifest; it also allows one to assert that the scientific view of the world recognises the need to recur to the sphere of persons – i.e. to the way they ordinarily conceive of themselves – in order to understand its genesis and justify its primacy.[15]

Without a genetic analysis of the scientific attitude in the historical human life-world, there is no possibility of understanding mathematics as a linguistic practice that constructs a world rather than merely describing it in a formalistic way. The opportunity a theory of the life-world offers is barred to an understanding that continues to think that the world is completely independent of the mind that thinks it. This last is, on the contrary, part of the physical system whose it can become a macroscopic observer. As we stressed, the product of our knowledge, i.e. what a scientist is entitled to call "reality," is always the result of an interaction between physical systems – in the case of the human mind enhanced by the mediation of technology and the elaboration of increasingly sophisticated languages. If we maintained that the phenomenon is what appears of something that remains hidden, we would not be thinking about the issue adequately. What we should claim is rather that reality is always only what appears on the basis of the way in which we interrogate it.

In the light of this, we can detect a metaphilosophical commitment within a theory of the life-world in a twofold sense.

Firstly, to question the need to resort to the sphere of persons even from a perspective that embraces scientific realism from the outset is to lay the foundations for a full justification of the latter. From this perspective, a theory of the life-world helps to clarify the sense of the primacy of the scientific image of the world.

[15] Let me notice that in science it is often inevitable to introduce descriptive terms that derive from the routine of the conceptual framework of persons in the manifest world: one speaks, indeed, of "corpuscles," "wave," "quantum foam," "scales," "black body," etc. As Sellars himself has pointed out, although it is necessary to make an effort to replace the observation language with theoretical language (and to understand that this does not imply that anything is left out), "only the most pythagoreanizing philosopher of science would attempt to dispense with descriptive (that is, nonlogical) predicates in his formulation of the scientific picture of the world" (1961: 126).

Secondly, more broadly but also more essentially, a philosophical enquiry into the life-world activates a philosophical reflection on what philosophy can do in its perennial attempt "to understand how things in the broadest possible sense of the term hang together in the broadest sense of the terms" (Sellars 1962: 1).

Danilo Manca
University of Pisa
danilo.manca@unipi.it

References

Argentieri, Niccolò, 2009, *Ci sono elettroni nel mondo-della-vita? Esperienza, matematica, realtà: una lettura fenomenologica dell'epistemologia di Werner Heisenberg*, Bonanno, Acireale-Roma.

Bell, John Stewart, 2004, *Speakable and Unspeakable in Quantum Mechanics*, Cambridge University Press, Cambridge.

Bernet, Rudolf, Kern, Iso, and Marbach, Eduard, 1989, *Edmund Husserl: Darstellung seines Denken*, Meiner Verlag, Hamburg; Eng. tr. with a foreword by Lester Embree, *An introduction to Husserlian Phenomenology*, Northwestern University Press, Evanston (IL), 1993.

Bitbol, Michel, 2020, "A Phenomenological Ontology for Physics: Merleau-Ponty and QBism", in Wiltsche, Berghofer 2020: 227-242.

Bohm, David, 1952, "A Suggested Interpretation of Quantum Theory in Terms of 'Hidden' Variables", in *Physical Review*, 85: 166-192.

Bohr, Niels 1981: *Collected Works*, general editor Léon Rosenfeld, vol. 2: *Work on Atomic Physics (1912-1917)*, ed. by Ulrich Hoyer, North-Holland Publishing Company, Amsterdam/New York/Oxford.

Born, Max, 1926, "Zur Quantenmechanik der Stossvorgänge", in *Zeitschrift für Physik*, 37: 863-867, eng. tr.: "On the quantum mechanics of collisions", in Wheeler, Zurek 1983: 52-61.

Born, Max, Heisenberg, Werner, and Jordan, Paul, 1926, "Zur Quantenmechanik II", in *Zeitschrift für Physik*, 35: 557-615; Eng. tr.: "On quantum mechanics II", in Van der Waerden 1967: 321-385.

Born, Max, and Jordan, Paul, 1925, "Zur Quantenmechanik", in *Zeitschrift für Physik*, 34: 858-888; partial Eng. tr.: "On quantum mechanics II", in Van der Waerden 1967: 277-306.

Cappelletti, Valentina, 2001, *Dall'ordine alle cose. Saggio su Werner Heisenberg*, Jaca Book, Milano.

Christias, Dionysis, 2020, "Lifeworld Phenomenology and Science", in Manca and Nunziante 2020: 261-286.

Christias, Dionysis, 2018, "Sellars' Naturalism, the Myth of the Given and Husserl's

Transcendental Phenomenology", *The Philosophical Forum* 49: 511-539.

De Broglie, Louis, 1925, "Recherches sur la théorie des Quanta", in *Annales de Physique*, 3/III (Janvrier-Février): 3-128.

De La Tremblaye, Laura, 2020, "QBism from a Phenomenological Point of View: Husserl and QBism", Wiltsche, Berghofer 2020: 243-260.

De Santis, Daniele, 2019, "Méditations Hégéliennes vs. Méditations Cartésiennes: Edmund Husserl and Wilfrid Sellars on the Given", in A. Ferrarin, D. Moran, E. Magrì, and D. Manca eds., *Hegel and Phenomenology*, Springer, Cham: 177-190.

Dirac, Paul, 1928, "The Quantum Theory of Electrons", in *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, 117/778 (Feb. 1): 610-624.

Einstein, Albert, 1931, "Maxwell's Influence on the Development of the Conception of Physical Reality", in James Clerk Maxwell, *A Commemoration Volume 1831-1931, with essays by Joseph John Thomson, Max Planck, Albert Einstein and others*, Cambridge University Press, Cambridge: 66-73.

Einstein, Albert, 1905, "Über einen die Erzeugung und Verwandlung des Lichtes betreffenden heuristischen Gesichtspunkt", in *Annalen der Physik*, 322/6: 132-148, eng. tr.: "Concerning an Heuristic Point of View Toward the Emission and Transformation of Light", in *American Journal of Physics*, 33/5, May 1965: 367-374.

Einstein, Albert, Born, Max and Hedwig, 1971, *The Born-Einstein Letters. Correspondence from 1916 to 1955*, with commentaries by M. Born, introduction by Werner Heinsenberg, foreword by Bertrand Russell, Eng. tr. Irene Born, Macmillan, London and Basingstoke.

Einstein, Albert, Podolsky, Boris, Rosen, Nathan 1935, "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?", in *Physical Review*, 47/10: 777-780.

Feynman, Richard P., 2010, *Lectures on Physics. The New Millennium Edition*, vol. I: *Mainly Mechanics, Radiation, and Heat*, Basic Books, New York.

Feynman, Richard P., 2006, *The Character of Physical Law*, Penguin Books, London.

Feynman, Richard P., 1983, *QED. The strange theory of light and matter*, Princeton University Press, Princeton (NJ).

French, 2002, "A phenomenological solution to the measurement problem? Husserl and the foundations of quantum mechanics", in *Studies in History and Philosophy of Modern Physics*, 33: 467-491.

French, 2020, "From a lost history to a New Future: Is a Phenomenological Approach to Quantum Physics viable?", in Wiltsche and Berghofer 2020: 205-224.

Greene, Brian, 1999, *The Elegant Universe: Superstrings, Hidden Dimensions, and the Quest for the Ultimate Theory*, W. W. Norton & Co. Inc., New York.

Heisenberg, Werner, 2000 [1958], *Physics and philosophy*, Penguin, London.

Heisenberg, Werner, 1994, *Ordnung der Wirklichkeit*, Piper Verlag, Münich; Eng. tr. by Martin B. Rumscheidt, Nancy Lukens and Irene Heisenberg, ed. by K.

Kleinknecht, 2019, *Reality and its Order*, Springer, Cham.

Heisenberg, Werner, 1927, "Über den anschaulichen Inhalt der quantentheoretischen Kinematik und Mechanik", in *Zeitschrift für Physik*, 43: 172-198; Eng. tr.: "The physical content of quantum kinematics and mechanics", in J.A. Wheeler and W.H. Zurek eds., 1983, *Quantum Theory and Measurement*, Princeton University Press, Princeton (NJ): 62-84.

Heisenberg, Werner, 1925, "Über quantentheoretische Umdeutung kinematischer und mechanischer Beziehungen", in *Zeitschrift fur Physik*, 33: 879-893; Eng. tr.: "Quantum-theoretical re-interpretation of kinematic and mechanical relations", in Van der Waerden 1967: 261-276.

Husserl, Edmund, 1968, "Umsturz der kopernikanischen Lehre", Ms. D17, first time published with the title "Grundlegende Untersuchungen zum phänomenologischen Ursprung der Räumlichkeit der Natur", in M. Farber, ed., *Philosophical Essays in Memory of Edmund Husserl*, Greenwood Press, New York: 307-325.

Husserl, Edmund, 1954, *Die Krisis der europäischen Wissenschaften und die transzendentale Phänomenologie. Eine Einleitung in die phänomenologische Philosophie*, Husserliana VI, edited by Walter Biemel, Nijhoff, Den Haag; Eng. tr. by David Carr, *The Crisis of European Sciences and Transcendental Phenomenology. An introduction to Phenomenological Philosophy*, Northwestern University Press, Evanston 1970.

Kerszberg, Pierre, 2012, *La science dans le monde de la vie*, Millon, Paris.

Lewis, Peter J., 2016, *Quantum Ontology. A Guide to the Metaphysics of Quantum Mechanics*, Oxford University Press, New York.

Lindley, David, 2007, *Uncertainty. Einstein, Heisenberg, Bohr, and the Struggle for the Soul of Science*, Anchor Books, New York.

Manca, Danilo, 2021, "Spontaneity and givenness: Natorp, Husserl, and Sellars's Neo-Kantianism", in *Phänomenologische Forschungen*, 2: 73-92.

Manca, Danilo, 2016, *Esperienza della ragione: Hegel e Husserl in dialogo*, Edizioni ETS, Pisa.

Manca, Danilo, 2014, "La scienza allo stato nascente: Merleau-Ponty e Sellars sull'immagine della natura", in *Discipline Filosofiche*, 24: 115-136.

Manca, Danilo, and Nunziante, Antonio M., eds., 2020, *Discipline Filosofiche*, vol. XXX, *Realismo, pragmatismo, naturalismo: Le trasformazioni della fenomenologia in Nord America*.

Maudlin, David, 2019, *Philosophy of physics. Quantum Theory*, Princeton University Press, Princeton and Oxford.

Maudlin, David, 2002, *Quantum Non-Locality and Relativity. Metaphysical Intimations of Modern Physics*, Blackwell, Oxford.

Maudlin, David, 1997, "Descrying the World in the Wave Function", in *The Monist*, 80/1: 3-23.

Rovelli, Carlo, 1996, "Relational Quantum Mechanics", in *International Journal of Theoretical Physics*, 35/8: 1637-1678.

Rovelli, Carlo, Laudisa, Federico, 2019, "Relational Quantum Mechanics", in *Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/entries/qm-relational/#Bib> (last reading October 1, 2021)

Sachs, Carl, 2020, "Phenomenology vs. the Myth of the Given: A Sellarsian Perspective on Husserl and Merleau-Ponty", in Manca and Nunziante 2020: 287-301.

Sachs, Carl, 2014, *Intentionality and the Myths of the Given: Between Pragmatism and Phenomenology*, Routledge, London.

Schrödinger, Erwin, 1928, *Collected Papers on Wave Mechanics*, Blackie & Son Limited, London and Glascow.

Schrödinger, Erwin, 1926a, "Quantisierung als Eigenwertproblem", in *Annalen der Physik*, 79/4: 361-376; 489-527; Eng. tr. "Quantisation as a Problem of Proper Values. Part I and II", in E. Schrödinger 1928: 1-40.

Schrödinger, Erwin, 1926b, "Über das Verhältnis der Heisenberg-Born-Jordanschen Quantenmechanik zu der meinen", in *Annalen der Physik*, 79/4: 734-756; Eng. tr. "On the relation between the Quantum mechanics of Heisenberg, Born, and Jordan", in E. Schrödinger 1928: 45-61.

Sellars, Wilfrid, 1991, *Science, perception and reality*, Ridgeview Company, Atascadero, 2nd edition.

Sellars, Wilfrid, 1981a, "Mental Events", in *Philosophical Studies: An International Journal for Philosophy in the Analytic Tradition*, 39/4: 325-345

Sellars, Wilfrid, 1981b, "Foundation for a Metaphysics of Pure Process", in *The Monist*, 64: 3-90.

Sellars, Wilfrid, 1968, *Science and Metaphysics*, Routledge & Kegan Paul, New York.

Sellars, Wilfrid, 1962, "Philosophy and the Scientific Image of Man", now in Sellars 1991: 1-40.

Sellars, Wilfrid, 1961, "The Language of Theories", now in Sellars 1991: 106-126.

Sellars, Wilfrid, 1956, "Empiricism and the Philosophy of Mind", now in Sellars 1991: 127-196.

Soffer, Gail, 2003, "Revisiting the Myth: Husserl and Sellars on the Given", in *The Review of Metaphysics*, 57: 301-337.

Trizio, Emiliano, 2021, *Philosophy's Nature: Husserl's Phenomenology, Natural Science, and Metaphysics*, Routledge, New York.

Van der Waerden, Bartelt Leendert, ed., 1967, *Sources of Quantum Mechanics*, Dover Publications, New York.

Van Fraassen, Bas C., 1980, *The Scientific Image*, Oxford University Press, New York.

Wheeler, John Archibald, Zurek, Wojciech Hubert, 1983, *Quantum Theory and Measurement*, Princeton University Press, Princeton (NJ).

Wiltsche, Harald A., Berghofer, 2020, *Phenomenological Approaches to Physics*, Springer, Cham.

Persons, Peirceish, perfidious pluralism – rescuing Sellars

Paul Giladi

Abstract: In *Philosophy and the Scientific Image of Man* (1962), Wilfrid Sellars contends that there is *tension* between manifest image (MI) and scientific image (SI) discursive formations. To end the tension and resolve the *clash* between the MI and the SI, Sellars does not aim to *reconcile* the two images. Rather, through the application of his functional classification semantics, typified by his distinction between logical irreducibility and causal reducibility, he aims to *join* the normative category of persons to the SI, to enrich and complete the SI. In other words, the way all things hang together stereoscopically in one unified and coherent image is by integrating persons into Peirceish. My principal aim in this paper is to argue that, rather than resolve the *clash* between the MI and the SI by joining the 'lifeworldy' conceptual framework of persons to the SI for the purpose of enriching and completing the SI, what Sellars ought to have done is adopt a *negative* dialectical 'resolution' of the clash between the images. This strategy invites one to dismantle the Placement Problem through the logic of "disintegration". I take Sellars to have curiously hinted at this Adornian intellectual orientation in the concluding sentence of *Empiricism and the Philosophy of Mind* (1956).

Keywords: Wilfrid Sellars; manifest image; scientific image; Peirceish; Theodor Adorno; negative dialectics.

1. *Fusion – not reconciliation: why this matters*

Dubbing Wilfrid Sellars 'a philosopher's philosopher' is worth repeating. Regardless of the ultimate appraisal of Sellars's account of the relationship between the manifest image of humanity-in-the-world (MI) and the Peirceish-regulated scientific image of humanity-in-the world (SI) in *Philosophy and the Scientific Image* (PSIM),[1] I think it is safe to say that Sellars's position is nothing if not *highly* nuanced. One evocative example of such nuance is the penulti-

[1] Peirceish is a discourse that construes *everything* in purely naturalistic descriptive terms for the purpose of constructive alignment with the causal explanatory interests of our best scientific theories.

philinq X, 1-2022, pp. 193-228 Submitted: January 2021
ISSN (print) 2281-8618, ETS doi: 10.4454/philinq.v10i1.386 Accepted: October 2021

mate sentence of PSIM, which can easily be underplayed or overlooked, by either those of a facile Hegelian disposition,[2] or some 'left-wing' Sellarsians, or some 'right-wing' Sellarsians:[3]

> The conceptual framework of persons is the framework in which we think of one another as sharing the community intentions which provide the ambience of principles and standards (above all, those which make meaningful discourse and rationality itself possible) within which we live our own individual lives. A person can almost be defined as a being that has intentions. Thus the conceptual framework of persons is not something that needs to be *reconciled with* the scientific image, but rather something to be *joined* to it. Thus, to complete the scientific image we need to enrich it *not* with more ways of saying what is the case, but with the language of community and individual intentions, so that by construing the actions we intend to do and the circumstances in which we intend to do them in scientific terms, we *directly* relate the world as conceived by scientific theory to our purposes, and make it *our* world and no longer an alien appendage to the world in which we do our living. (SPR: 40)

Prima facie, the difference between (a) reconciling the 'lifeworldy' conceptual framework of persons and the language of community and individual intentions, with the categorial ontology and discursive formations of the SI, and (b) joining the 'lifeworldy' conceptual framework of persons to the SI for the purpose of enriching and completing the SI, seems stylistically, not substantively, different. This is because both the act of reconciliation and the act of joining result (if successful) in ending the clash between the images, which is what Sellars principally aims to accomplish.

However, just because the respective acts of 'reconciling' and 'joining' (if successful) end the conflict between 'the perennial philosophy' and 'postulational scientific strategy,' *this* in and of itself provides no legitimate reason to regard 'the reconciliation of the MI with the SI' and 'the joining of the MI to the SI' as semantically interchangeable. For, Sellars makes it clear in the quoted passage above that he envisions a *fusion* between the two, rather than a reconciliation between the two discursive drives and cognitive orientations.[4] The conceptual framework of persons and the language of community

[2] Sellars frustratingly construes the synoptic vision in terms of "synthesis" (SPR: 9). I will return to this issue in the latter stages of the paper, starting with IIIa.

[3] Left-wing Sellarsians (most notably Richard Rorty, Quill Kukla, Robert Brandom, John McDowell, and Michael Williams) prioritise Sellars's critique of the Myth of the Given. Right-wing Sellarsians (most notably Ruth Millikan, Patricia Churchland, Paul Churchland, William Lycan, Jay Rosenberg, Daniel Dennett, and Johanna Seibt) prioritise Sellars's commitment to a strong form of scientific realism and his commitment to a prescriptive variety of naturalism at the ontological level.

[4] However, what complicates the substantive (as opposed to merely verbal) distinction between

and individual intentions are to be *integrated* with the conceptual framework of postulational science and the "doggedly naturalistic" (O'Shea 2007: 187) language of description and explanation for a specific purpose and epistemic achievement: to enrich and complete the SI. As Dionysis Christias puts it, the idea here is "a smooth incorporation of normativity within the scientific image" (Christias 2016a: 460).

Sellars is not disposed to 'reconcile' the MI and the SI if that were to mean assuming (i) that they are *equal* rivals in *all* functional forms of discursive life (in *both* representational [descriptive-explanatory] and expressive [normative] discourse), and (ii) that they are mutually dependent for each other's enrichment and completeness.[5] Rather, Sellars is disposed to finding a way to only enrich and complete the SI by adding the MI's normative-centric vocabulary to the SI. Crucially, the MI is not to be enriched and completed – it will, in an important sense, 'wither away' as the SI develops. Talk of 'reconciling,' not that of 'joining' – *contra* Willem deVries (2016b: 119) – is at best misleading for appreciating the details of Sellars's "stereoscopic vision" (SPR: 9). The difference between (a) and (b), therefore, is *substantively* different.

2.

2.1. Irreducible discontinuity

As is well known, Sellars structures his project in PSIM around the aim of revealing

> how things in the broadest possible sense of the term hang together in the broadest possible sense of the term … To achieve success in philosophy would be [...] to 'know one's way around' with respect to all these things [...] in that reflective way which means that no intellectual holds are barred. (SPR: 1)

'joining *x* to *y*' and 'reconciling *x* with *y*,' to the extent that Sellars's argument concerns how to join the MI to the SI (rather than reconcile the two), is the following passage from PSIM: "… the task of showing that categories pertaining to man as a *person* who finds himself confronted by standards (ethical, logical, etc.) which often conflict with his desires and impulses, and to which he may or may not confirm, can be reconciled with the idea that man is what science says he is" (SPR: 38).

[5] Jay Rosenberg appears to switch between fusion-talk and reconciliation-talk as if there is no *substantive* difference between the two *types* of talk: "Much of Sellars's philosophical work can be understood as an attempt to show how the person-constitutive categories of the manifest image might be reconciled with or coherently added to the scientific image to produce a 'stereoscopic' or 'synoptic' image" (J. Rosenberg 2009: 285).

In other words, Sellars sees the function of the best philosophical inquiry not to rest content with the first-order task of carving reality at its joints, but to engage with the second-order task of doing justice to the inherently polydimensional structure of various logical spaces. Sellars insists that banal natural kinds such as cabbages, as well as scientifically "recalcitrant" (Rosenthal 2016: 150), 'queerer'[6] kinds, such as duties and aesthetic experience, are made sense of in equal measure (cf. SPR: 4). In doing so, he contrasts the "the analytic conception of philosophy as myopia" – namely, the first-order task of carving reality at its joints – with "the synoptic vision of true philosophy" (SPR: 3), his second-order task of doing justice to polydimensional discourse.

Significantly, in PSIM, the reflectively-minded philosopher is confronted by an antimony:

> I want to highlight from the very beginning what might be called the paradox of man's encounter with himself, the paradox consisting of the fact that man couldn't be man until he encountered himself […] Its central theme is the idea that anything which can properly be called conceptual thinking can occur only within a framework of conceptual thinking in terms of which it can be criticised, supported, refuted, in short, evaluated […] The attempt to understand this […] turns out to be part and parcel of the attempt to encompass in one view the two images of man-in-the-world which I have set out to describe. For, as we shall see, this difference in level appears as an irreducible discontinuity in the *manifest* image, but as, in a sense requiring careful analysis, a reducible difference in the *scientific* image. (SPR: 6)

Sellars's claim that there is "an irreducible discontinuity in the *manifest* image, but as, in a sense requiring careful analysis, a reducible difference in the *scientific* image" serves as his nuanced strategy for resolving the antinomy by "stereoscopic vision, where two differing perspectives on a landscape are fused into one coherent experience" (SPR: 4). To quote Stephanie Dach here, "the process is one where the conceptual frameworks of the MI and the SI need to be mutually adapted" (Dach 2018: 572; cf. Christias 2019b: 464), to the extent that the mutual adaptation here – which is necessary for integration/fusion by stereoscopic envisioning – involves a specific sense in which the MI is irreducible to the SI, and a different, incommensurable sense in which the MI is reducible to the SI.

In what immediately follows, I will explain, following Jay Rosenberg (2007a; 2007b), Christias (2016a; 2016b; 2019b) and James O'Shea (2007; 2009; 2016), that the stereoscopic vision is best made sense of by situating it against

[6] See Giladi 2019a, 2020a for further on the characterisation of normative kinds as 'queer' or scientifically recalcitrant.

the backdrop of Sellars's "functional classification" semantics (J. Rosenberg 2007a: 15), typified by his 1953 distinction between logical irreducibility and causal reducibility.

For Sellars, the conceptual framework of persons, the hallmark of the MI, is *logically* irreducible to the discursive formations and categorial framework of the SI. What this means is that the language and 'constitutive-interest,' of personhood is incapable of translation into the language and framework of natural science (see Habermas 1971: 196). As Sellars writes:

> [w]hatever users of normative discourse may be conveying about themselves and their community when they use normative discourse, what they are saying cannot be said without using normative discourse. The task of the philosopher cannot be to show how, in principle, what is said by normative discourse could be said without normative discourse, for the simple reason that this cannot be done. His task is rather to exhibit the complex relationships which exist between normative and other modes of discourse. (PPPW: 82)

> [O]ne scarcely needs to point out these days that however intimately conceptual thinking is related to sensations and images, it cannot be equated with them, nor with complexes consisting of them. (SPR: 32)

> Indeed, there are sound methodological reasons for not teaching ourselves to respond to perceptual situations in terms of constructs in the language of theoretical physics. For while this could, in principle, be done, the scientific quest is not yet over, and even granting that the main outlines are blocked in, the framework of physical objects in space and time, shaped over millennia of social evolution, provides, when accompanied by correct philosophical commentary, a firm base of operations with which to correlate the developing structure of scientific theory, refusing to embrace any stage without reverse as our very way of perceiving the world, *not* because it wouldn't be a *better* way, but because the better is the enemy of the best (SPR: 97).

> Now the idea that epistemic facts can be analysed without remainder – even 'in principle' – into non-epistemic facts … with no matter how lavish a sprinkling of subjunctives and hypotheticals is, I believe, a radical mistake – a mistake of a piece with the so-called 'naturalistic fallacy' in ethics (SPR: 131).

If one aims to either logically reduce sentences involving first-person intentional vocabulary to sentences involving purely non-intentional vocabularies (cf. Fodor 1974: 104; Wedgwood 2007: 145), or to even Ramsify sentences involving first-person intentional vocabulary (viz. Ramsey 1931; Lewis 1970;

Jackson 1998: 140), then such semantic tasks involve translating the *intentional* MI framework into the *extensional* SI framework.[7]

Focusing on Ramsification at least, suppose T refers to the theoretical terms of first-person psychology; and suppose that $O_1 \ldots O_n$ stands for a set of heterophenomenological predicates, such as 'is a concurrent neurophysiological event' and 'has minimising prediction error.' To Ramsify T, then, one replaces the first-person psychological terms with the existentially bound variables of relevant empirical terms (i.e. observable heterophenomenological descriptors): $R(T) = (\exists!)O_1 \ldots O_n\, T[O_1 \ldots O_n]$. In the Ramsey Sentence '$R(T)$,' theoretical terms are logically eliminated, leaving only the empirical content of the observational sentence as the semantic content of $R(T)$ (viz. Carnap 1975: 82-83). In the case of Ramsifying first-person psychology, normative content does not feature as part of the semantic content of $R(T)$, whose predicates are only cognitive neuroscientific descriptors.

However, this effort to bring about a cognitively meaningful language shorn of intentional content is ultimately doomed to failure for at least two principal reasons. First, normative vocabulary underpins the logic of Ramsification and its revisionary extensional semantics in the first place, as epistemic virtues, such as plausibility and simplicity, motivate the very exercise of swapping the theoretical terms with the bound variables of regimented observational sentences (see Putnam 2002: 30-31; 141). Second, Ramsification – as well as logical reductionism – fails to make cogent enough *functional sense* of sentences that have ostensible commitments to "iffy" normative kinds (SPR: 24; cf. Baker 2013: 35; xv). As Sellars writes regarding the latter point, "[…] such a reconstruction [translating the intentional MI framework into the extensional SI framework] is *in principle* impossible, the impossibility in question being a strictly logical one" (SPR: 38).

Crucially, it is important to note how the following two positions, despite sharing much in common with Sellars's position, are *not* involved in how he steers his argument that the conceptual framework of persons is logically irreducible to the SI. First, Sellars's argument does not principally make use of a general anti-positivist[8] commitment to the autonomy of philosophy in the face of the colonisation of MI spaces and categories by devout followers of an unqualified, imperialistic, hierarchical Unity of Science Thesis (UIHUST).[9]

[7] I recognise that it is a widely and seemingly unquestioned assumption that the complete SI would be described by a purely extensional semantics.

[8] This is not to say that Sellars was devoid of anti-positivist commitments. For example, SPR: 20-22 evinces Sellars's *qualified* Unity of Science Thesis. As Michael Hicks notes, "[t]o see Sellars's ontological monism as requiring of him a naive conception of the unity of science is to accuse him of precisely the view he means to be criticising" (Hicks: forthcoming). Cf. O'Shea 2007: 45.

[9] See Nagel 1961; Oppenheim and Putnam 1958; A. Rosenberg 2014.

Such a position contends that every phenomenon explicable by special sciences, such as biology and psychology, is in principle reductively explicable by fundamental physics. Suffice to say that UIHUST is naïve, 'greedy' (in Daniel Dennett's sense), and, above all, *easily* refutable, so much so that UIHUST is not taken especially seriously in the philosophy of science.[10] Second, Sellars's argument for the logical irreducibility of normativity does not make use of either (i) P.F. Strawson's Kantian commitment to how "the absence of the reactive attitudes is a world of human isolation so cold and dreary that any but the most cynical must shudder at the idea of it" (Wolf 2008: 73) or (ii) Donald Davidson's anomalism about the mental.

Regarding (i), Strawson's descriptive metaphysical defence of reactive attitudes as core features of our ordinary conceptual scheme is not a strategy that plays a role in Sellars's argument. For that matter, if anything, there is more compelling reason to think of Sellarsian metaphysics as a hybrid, *sui generis revisionary* and *descriptive* project, rather than in terms of a straightforwardly descriptive focus on the core features of our actual conceptual scheme: Sellars tries to combine descriptive metaphysical interests about ordinary language and our everyday practices (à la Strawson, J.L. Austin, and C.I. Lewis) with a revisionary metaphysics of science (the process philosophy of A.N. Whitehead), and dialectically play them off each other until they settle into a mutually supporting structure.

On the subject of (ii), Davidson's commitment to the nomological (as opposed to logical) irreducibility of mental predicates *subtly* differs from Sellars's strategy. According to Davidson, "there may well exist a physical open sentence coextensive with each mental predicate" (Davidson 2001: 215-216). However, though Davidson qualifies this possibility and avoids clear Ramsification, writing that "to construct [a physical open sentence coextensive with each mental predicate] might involve the tedium of a lengthy and uninstructive alternation" (Davidson 2001: 216), Davidson's position is different to Sellars's. This is because, for Sellars, logical irreducibility is not a failure of coextension: it is a failure of analyticity. Sellars, unlike Davidson, does *not* follow Quine in rejecting the analytic/synthetic distinction. Davidson accepts that semantics must be extensional. As such, he is committed to the idea that the coextensive physical open sentence may be so long that it becomes unworkable. Sellars, however, following Rudolf Carnap and Lewis, never abandoned *intensional* semantics. For him, the whole point of 'meaning as functional classification' is

[10] I think it is worth emphasising that, as Cartwright *et al.* 1996 has convincingly argued, UIHUST is *not* attributable to Otto Neurath, especially considering Neurath's anti-foundationalism, anti-pyramidism, and articulation of an 'encyclopaedia-model.'

to make intensions compatible with a prescriptive naturalistic ontology.

The driving force of Sellars's argument is his use of functional classification semantics to elaborate the naturalistic fallacy:

> Now the idea that epistemic facts can be analysed without remainder – even 'in principle' – into non-epistemic facts […] with no matter how lavish a sprinkling of subjunctives and hypotheticals is, I believe, a radical mistake – a mistake of a piece with the so-called 'naturalistic fallacy' in ethics. (SPR: 131)

> To say that a certain person desired to do *A*, thought it his duty to do *B* but was forced to do *C*, is not to describe him as one might *describe* a scientific specimen. One does, indeed, describe him, but one does something more. And it is this something more which is the irreducible core framework of persons […] Now, the fundamental principles of a community, which define what is 'correct' or 'incorrect,' 'right' or 'wrong,' 'done' or 'not done,' are the most general common *intentions* of that community with respect to the behaviour of the members of the group. It follows that to recognise a featherless biped or dolphin or Martian as a person requires that one think thoughts of the form 'We (one) shall do (or abstain from doing) actions of kind *A* in circumstances of kind *C*.' To think thoughts of this kind is not to *classify* or *explain*, but to *rehearse an intention* (SPR: 39-40).

Making sense of persons (and their entangled normative categories, such as agency and knowledge) in *this* cognitive context functionally means that our sense-making framework is not exclusively in the business of describing and explaining. The cognitive context of the specific discourse concerning persons and their associated normative categories is one typified by *expressive* as opposed to *representational* speech-acts.[11] In Hegelian-pragmatist fashion, Sellars holds that individuating persons is not determined by a description of person-practices, but rather by an account of how *these* specific practices convey persons' sensitivity to a normative community (cf. Levine 2019: 253), the ways in which persons are sensitive to fellow language-using, norm-bearing agents (cf. Kukla and Lance 2009: 185), the ways in which persons occupy a recognisable standing in the social space of reasons. As Quill Kukla (writing as Rebecca Kukla) and Mark Lance point out,

> Sellars is getting at the point that recognising someone as a person is not merely an observative act, but also a practical act of the second kind … We become and remain

[11] Cf. "[…] [W]e cannot intelligibly attribute any propositional attitude to an agent except within the framework of a viable theory of his beliefs, desires, intentions, and decisions … [W]e make sense of particular beliefs only as they cohere with other beliefs, with preferences, with intentions, hopes, fears, expectations, and the rest" (Davidson 2001: 221).

the types of beings that have specific, agent-relative engagements with others through an ongoing network of hails and acknowledgments. (Kukla and Lance 2009: 180-181)

The discursive morphology of the conceptual framework of persons regards representational discourse as *functionally* unsuitable for facing up to and making sense of normative kinds *qua* normative kinds. This is because making sense of these 'iffy' and scientifically 'recalcitrant' phenomena is the business – or constitutive-interest – of *expressive* discourse: persons are not 'emergent' kinds over and above the descriptive-explanatory categories of science; the categorial status of persons is that of an entity with a recognisable standing in the logical space of reasons, rather than a non-aggregative natural kind.[12] Understood in this way, while the emergentist discourse of non-aggregativity is anti-reductionist, it is anti-reductionist in a functionally different way to Sellars's notion of logical irreducibility, insofar as non-aggregativity is a *representationalist* species of anti-reductionism. Talk of recognisable standings in the logical space of reasons, since this talk is *expressive*, is not of interest to any descriptive and explanatory projects in the first place, given their functional role differentiation. As Richard Bernstein writes,

> [e]verything that can be described and explained about persons can be described and explained in terms of the scientific image. The 'something more' that is left over is not something more to be described and explained; it involves the having and rehearsing of intention. (1966: 125)

To repeat Sellars's oft-quoted passage in *Empiricism and the Philosophy of Mind* (EPM), "[on the subject of normative kinds (such as knowledge, meaning, and persons),] we are not giving an empirical description … we are placing [them] in the logical space of reasons, of justifying and being able to justify what one says" (SPR: 169).

This irreducible discontinuity in the MI – logical irreducibility – is only half the Sellarsian story. To complete the Sellarsian story, one needs to articulate what he means by 'causal reducibility' – the reducible difference in the SI.

2.2. Reducible difference

Sellars defines 'causal reducibility' in a way that is bound up with (i) his *scientia mensura* principle in EPM, and (ii) his particular take on the error underpinning the naturalistic fallacy:

[12] Hegel, *contra* Sellars, regards persons as emergent kinds: "Spirit has thus proceeded from Nature […] But it is one-sided to regard spirit in this way as having only *become* an actual existence after being merely a potentiality …" (EPW §376Z; Eng. tr. 1970: 444).

… in the dimension of describing and explaining the world, science is the measure of all things, of what is that it is, and of what is not that is not. (SPR: 173)

If we use 'ethical assertion' in such a way that 'Jones ought to pay his debt' is an ethical assertion, but 'Jones feels that he ought to pay his debt' is not, then we can say that to claim that Ought is causally reducible to Is is to claim that one can give a causal explanation of the history of moral agents without making ethical assertions (PPPW: 48-9).

[A] concept will be said to be causally reducible to descriptive concepts if (roughly) it [...] occurs in the antecedent of a properly constructed casual explanation only as a subordinate element in a descriptive mentalistic context [...] Thus, a non-naturalist who holds that the only way in which moral obligation can enter into the causal explanation of human history is via facts of the form Jones thinks (feels) that he ought to pay his debt, would be holding that Ought is, in the above sense, causally reducible to Is. In traditional terminology, he would be claiming that obligation enters into the causal order only as an element in the intentional object of a mental act. (PPPW: 122)

Sellars distinguishes causal reducibility from logical reducibility in terms of how causal reducibility is concerned with descriptive-explanatory power, which is functionally "orthogonal" to logical irreducibility (Levine 2019: 255). The primacy of the SI consists in how the SI, rather than the MI, instructs on ontological matters.[13] Understood in this manner, the MI will 'wither away' *on the ontological side of sense-making* (cf. O'Shea 2007: 162), since in the domains of description and explanation, the SI is better at finding out what there is than sophisticated common sense.[14] Significantly, the descriptive-explanatory function of *causally* reducing persons to "a complex physical system" (SPR: 22) is to make sense of persons *qua* a postulational SI category devoid of *any* commitment to (the use of) normative terms. In other words, a causal reduction of 'person' to 'a complex physical system' is the representational function of Peirceish:

[13] Viz. "But, *speaking as a philosopher*, I am quite prepared to say that the common sense world of physical objects in Space and Time is unreal – that is, that there are no such things" (SPR: 173). Viz. "The framework of common sense is radically false (i.e. there *really* are no such things as the physical objects and processes of the common sense framework)" (PP: 354).

[14] Quine (1948) construes ontological commitment in terms of being a bound variable of a regimented sentence of our best current science. Sellars's metalinguistic functional role criterion for ontological commitment, however, "*explains* the syncategorematic character of predicates *without any reference to quantification*" (N&O: 51). Sellars insists that our best current science tells us what there is because our best current science causally "hook[s] up" (N&O: 10) with the measurable, determinate objects explicitly named in true empirical propositions. Determinate reference, not indeterminate reference, explains "*how* 'variables of quantification' hook up with the world" (N&O: 10). Determinate reference, not indeterminate reference, explains *how* true empirical sentences 'picture' objects.

[t]he naturalistic 'thesis' that the world, including the verbal behaviour of those who use the term 'ought' – and the mental states involving the concept to which this world gives expression – can 'in principle,' be described without using the term 'ought' or any prescriptive expression, is a logical point about what it is to count as a description *in principle* of the world … [N]aturalism presents us with the ideal of a *pure* description of the world (in particular human behaviour), a description which simply says what things *are*, and never, in any respect, what they *ought* or *ought not* to be; and it is clear (as a matter of simple logic) that neither 'ought' nor any other prescriptive expression *could be used* (as opposed to *mentioned*) in such a description. (CDCM: 283)

To quote deVries (2019b: 233) here, "[t]he description is not 'gappy' at the level of the fundamental ontology of the natural world". Through the application of his functional classification semantics, typified by his distinction between logical irreducibility and causal reducibility, Sellars aspires to uphold the primacy of the SI "without reaching for … an eliminativist sledgehammer" (O'Shea 2009: 194)[15] In other words, the way all things hang together stereoscopically in one unified and coherent image is by integrating persons into the Peirceish-regulated model of describing and explaining.

I think it is reasonable to claim that Sellars's commitment to anti-foundationalism and expansive conceptual frameworks, typified by his notion of synoptic vision, aims to be democratic and non-supremacist, because his position makes it clear, given his frequent use of 'polydimensional,' that there must be a pluralism of vocabularies in play to adequately make sense of things. Above all, Sellars's commitment to anti-foundationalism and to the synoptic vision paints a reasonably convincing picture of a thinker who wishes to replace, as Adriana Cavarero (2016) would phrase it, a *rectitudinal*, fixed, and vertical image of sense-making practice with a stereoscopic, dynamic, and horizontal image of sense-making practice. Such commitments *eo ipso* do not involve a commitment to (1) UIHUST and/or (2) a conceptually crude and crass variety of naturalism (cf. deVries 2019a: 37).

Thus far, I have reconstructed how Sellars envisions the joining of the conceptual framework of persons to Peirceish. Like O'Shea and Christias, I think Sellars's stereoscopic vision is best construed as a functionalist 'naturalism with a normative turn,' in that persons are logically irreducible but causally reducible to the descriptive-explanatory categories of science (cf. O'Shea 2009: 207). This is what I take Sellars's position to be. The functional classification distinction between normative discourse and descriptive-explanatory dis-

[15] Cf. deVries 2017: 1647: "Sellars never intimates that future science will give us occasion to discard folk psychology, although some of Sellars's students have drawn that conclusion". See also Christias 2016b: 2854.

course not only anchors the important Kant-inspired Sellarsian claim that "[t]he scientific image and the framework of persons … have different tasks that operate in different domains" (Levine 2019: 255). In what immediately follows, I will argue that there is a danger that Sellars's idiosyncratic *Aufhebung* of persons *nonetheless* risks erasing persons and the MI's discursive form of life.

3.

3.1. The "Eye on the whole" and Begriffsbildung – the opening Adornian salvo

In 1997, Richard Rorty (1997: 8-9) remarked that Sellars's project was "an attempt to usher analytic philosophy out of its Humean and into its Kantian stage;" whereas "[Robert] Brandom's project can usefully be seen as an attempt to usher analytic philosophy from its Kantian to its Hegelian stage". Whatever truth there may be in this characterisation by Rorty, it nevertheless underestimates how deeply Hegelian Sellars himself already was.[16] On this specific subject, deVries (2017: 1653) has argued that

> Hegel and Sellars are both, in the end, monistic visionaries who try to explicate how it is possible for finite subjectivities to grasp the reality around them as it is in itself. No distinctions are primitive givens for them; each distinction must be justified, for, in the end, the world is One.

Sellars, much like Hegel before him, has an "'eye on the whole'" (SPR: 3). However, symptomatic of Western metaphysics, according to Theodor Adorno, is 'identity philosophy' – which represents the long-standing pathological cognitive propensity, beginning with Plato and 'actualised' by Hegel, to prioritise universality over individuality. For Adorno, "[i]dentity is the primal form of ideology" (1973: 148). The prioritisation of universality (and the concomitant concepts of unity and identity) is regarded by Adorno as harmful, insofar as the underpinning practice of conceptualisation (*Begriffsbildung*) here, namely the discursive operations of the 'eye on the whole,' is inherently violent and authoritarian. This is because, for Adorno, non-identity[17] and difference are

[16] Cf. deVries 2017: 1648: "however much of the Hegelian wine Sellars preserves, he is persistent in re-bottling it in naturalistic flasks". For detailed discussions about the relationship between Hegel and Sellars, see issue 3 of Volume 27 of the *International Journal of Philosophical Studies*.

[17] Viz. Adorno 2008: 7: "[B]y subsuming them all under this concept, by saying that A is everything that is comprehended in this unity, I necessarily include countless characteristics that are not integrated into the individual elements contained in this concept. The concept is always less than what is subsumed under it. When a B is defined as an A, it is always also different from and more than the A, the concept under which it is subsumed by way of a predicative judgement. On the other hand,

invariably sacrificed on the altar of unity-in-the-system.[18] Parts are nothing; the whole is everything. To quote Espen Hammer (2020: 41) here, "[f]or Adorno, our concepts do what King Midas did when his wish for ever more gold was granted him – they turn what's *living and different*, yet potentially intimate, into a *dead, repetitive sameness*".

Since *Begriffe* function to *seize* the things at which they are directed,[19] the cognitive activity of making sense of things through the application of totalising rule-conforming concepts is coercive, and does not respect the diverse integrity of existence itself. Rather, if anything, the 'eye on the whole' and *Begriffsbildung*, for Adorno, are effectively a kind of *viol cognitif*, where reality is brutally forced to conform to totalising discursive categories. Such cognitive totalising risks translating into a form of *social* totalitarianism, because the activity of subsuming under a whole leads to assimilating "all individuals into a general type, and thereby exclude or devalue their difference or singularity" (Stern 2009: 367). As Brian O'Connor (2013: 82) notes, "… for Adorno, this form of coercion is precisely what happens at the level of modern social organisation. This is no coincidence". These points are vividly expressed by Adorno in the two passages below:

> unity gets worse as its seizure of plurality becomes more thorough. It has its praise bestowed on it by the victor, and even a spiritual victor will not do without his triumphal parade, without the ostentatious pretence that what is incessantly inflicted upon the many is the meaning of the world […] Thus established, the logical primacy of the universal provides a fundament for the social and political primacy that Hegel is opting for. (1973: 328)

> The conception of a totality harmonious through all its antagonisms compels [Hegel] to assign to individuation, however much he may designate it a driving moment in the process, an inferior status in the construction of the whole … [W]ith serene indifference [Hegel] opts once again for liquidation of the particular. Nowhere in his work is the primacy of the whole doubted (1974: 16-17).

To add textual weight to Adorno's caustic critique of Hegel, one might point to the following passages from Hegel's *Encyclopaedia Logic*, where the

however, in a sense every concept is at the same time more than the characteristics that are subsumed under it".

[18] Viz.: "What tolerates nothing that is not like itself thwarts the reconcilement for which it mistakes itself. The violence of equality-mongering reproduces the contradiction it eliminates" (Adorno 1973: 142-143).

[19] The German term for 'concept,' *Begriff*, comes from the verb *Begreifen*, which in turn is derived from *Greifen*. 'Greifen' is often translated as meaning 'to grab' / 'to grip' / 'to seize' / 'to snatch' / 'to capture' / 'to strike' / 'to take hold' / 'to bite.'

operation of reason (*Vernunft*) seems identical to the (aggressive) operation of the understanding (*Verstand*):[20]

> [w]e feel the need to bring unity to this manifold; therefore, we compare them and seek to [re]cognise what is universal in each of them. Individuals are born and pass away; in them their kind is what abides, what recurs in all of them; and it is only present for us when we think about them … in thinking about things, we always seek what is fixed, persisting, and inwardly determined, and what governs the particular (EL: §21Z, 53; emphasis added).

> What human beings strive for in general is cognition of the world; we strive to appropriate it and to conquer it. To this end the reality of the world must be crushed as it were; i.e., it must be made ideal. At the same time, however, it must be remarked that it is not the subjective activity of self-consciousness that introduces absolute unity into the multiplicity in question; rather, this identity is the Absolute, genuineness itself. Thus it is the goodness of the Absolute, so to speak, that lets singular [beings] enjoy their own selves, and it is just this that drives them back into absolute unity (EL: §42, 85 – emphasis added).

The presence of "govern" is especially important here, as this precisely seems to motivate (i) Adorno's specific logico-metaphysical polemic that the logic of domination operates in the discursive operations of the 'eye on the whole,' and (ii) Adorno's 'negative dialectic' reversal of Hegel's claim in the *Phenomenology of Spirit* (§20, 13) that *das Ganze ist das Wahre*:[21]

> "The whole is the untrue," not merely because the thesis of totality is itself untruth, being the principle of domination inflated to the absolute; the idea of a positivity that can master everything that opposes it through the superior power of a comprehending spirit is the mirror image of the experience of the superior coercive force inherent in everything that exists by virtue of its consolidation under domination. (Adorno 1993 [1963]: 87)

According to Adorno, the logical structure of modern social organisation is typified by drives towards the domination (and even obliteration) of differ-

[20] As Todd McGowan phrases it, "[t]he understanding is the vehicle of epistemic violence" (2019: 73).

[21] Viz. Adorno 2008: 18: "Now, when I speak of 'negative dialectics' not the least important reason for doing so is my desire to dissociate myself from this fetishisation of the positive, particularly since I know full well that the concept has an ideological resonance that is connected with the advances made by certain philosophical trends and that very few people are aware of". Viz. 2008: 20: "Unlike the kind of dialectics that the late Hegel called for, one in which the affirmative could be discovered at the end of all the negations, this concept calls for the very opposite".

ence: "we are dealing with the principle of mastery" (2008: 9). These steering mechanisms geared towards universal reification produce a 'false totality,' to use Max Horkheimer's term, and – at the material-psychological level – result in a damaged subjectivity, damaged life. The function of negative dialectics, therefore, is not to offer a liberal-inspired *resistance* to these totalising dispositions and ideological forms of modern social organisation. Rather, the function of negative dialectics is radical, construed as a *reversal* of logico-metaphysical power, according to which the category of difference (namely, non-identity) has *priority* over totalising categorial frameworks.[22] As Adorno himself writes on this subject, "[t]o change this direction of conceptuality, to give it a turn toward non-identity, is the hinge of negative dialectics" (1973: 12).

The 'eye on the whole,' as such, invariably becomes an oppressive gaze, since "the unity of the self-preserving thought [as a "rationalised rage at non-identity"] may devour it without misgivings" (1973: 23). As one instantiation of the logical structure of modern social organisation geared towards increasing homogenisation, the 'eye on the whole' is guilty of a cognitive variety of imperialism, where such a position is the theoretical equivalent of Iris Marion Young's concept of cultural imperialism:[23]

> [i]n societies stamped with cultural imperialism, groups suffering from this form of oppression stand in a paradoxical position. They are understood in terms of crude stereotypes that do not accurately portray individual group members but also assume a mask of invisibility; they are both badly misrepresented and robbed of the means by which to express their perspective. Groups who live with cultural imperialism find themselves defined externally, positioned by a web of meanings that arise elsewhere. These meanings and definitions have been imposed on them by people who cannot identify with them and with whom they cannot identify. (Young 1990: 59)

For Young, most modern societies contain multiple cultural groups, some of which unjustly dominate the state or other important social institutions, thus inhibiting the ability of minority cultures to live fully meaningful lives in *their own terms*. The dominant group in society can limit the ability of one or

[22] Cf. Hammer 2020: 37: "… the exercise of "negative dialectics" – the attempt to reveal, rather than overcome, the *dis-unifications* … that contemporary social practice keeps effacing. It becomes, one might say, a form a radical ideology critique whereby our various modalities and practices of identification, sense-making, and conceptualisation, including those of academic philosophy, are subjected to critique". Cf. O'Connor 2013: 102.

[23] In the 1990s, the politics of difference focused on questions concerning nationality, ethnicity, and religion. Under this approach, the value of cultural distinctness is *essential* to individuals and not something accidental to them: their personal autonomy depends in part on being able to engage in specific cultural practices with others who identify with one another as in the same cultural group.

more of the cultural minorities to live out their forms of expression. In other words, the dominant culture threatens to swamp the minority culture, to the extent that particular cultural practices and different hermeneutic spheres – ways in which members of cultures interpret their experiences – are crowded out or erased.

How does this bear on the question about Sellarsian naturalism, though? I contend that the Adornian concern about Sellars's naturalism-with-a-normative-turn is that "[u]ltimately the scientific image will drown out the manifest image" (Hicks: forthcoming): the 'eye on the whole' and *Aufhebung* of persons, regardless of any left-wing Sellarsian intentions, threatens to swamp persons and the language of community and individual intentions (the MI's discursive form of life). The vocabulary of Peirceish is epistemically authoritarian, in that it really risks forcing other forms of inquiry to adopt the discursive recourses and grammars of formal disciplines categorially different to the MI's discursive and grammatical constellations (cf. McDowell 1994: 70):

> [s]cientific objectification, in line with the quantifying tendency of all science since Descartes, tends to eliminate qualities and to transform them into measurable definitions. Increasingly, rationality itself is equated *more mathematico* with the faculty of quantification. While perfectly corresponding to the primacy of a triumphant natural science. (1973: 43)

Because the Sellarsian synoptic vision is *primarily* structured by the doggedly sparse physicalist ontology of the SI, the purely naturalistic vocabulary will invariably fail to fulfil the function of mitigating conflict with the constellation of persons and the language of community and individual intentions, since the conceptual framework of persons is subject to regulatory discourse, insofar as they must be *forced* into naturalistic categories that does not seem appropriate for *their specific* characters. Naturalistic categories and empirical science itself are the products of the domination, not revelation, of nature. Therefore, the doggedly sparse physicalist ontology of the SI is not as innocent as Sellars makes it out to be.

Under the synoptic vision, there is little or no way to epistemically counter colonisation and eventual obliteration by the SI, since what is the *base* of the synoptic vision *superstructure* is purely naturalistic vocabulary. If the base is constituted by Peirceish, then securing and protecting persons as agents and as conceptual thinkers within the doggedly sparse physicalist ontology of the SI is effectively impossible (see O'Shea 2009: 194).

3.2. The (left-wing) Sellarsian response

I think it is reasonable to contend here that Sellars would be rather unimpressed, and perhaps even cholerically frustrated, with this Adornian critique. For, the very idea of using the base-superstructure framework as a way of critically making sense of the logical architecture of stereoscopic thinking not only strangely charges Sellars with foundationalism, it also categorically misrepresents, to the point of even ignoring, Sellars's nuanced notion of unity bound up with his functional classification semantics. In what follows, I construct a Sellarsian rebuttal to my opening Adornian salvo.

In a curiously Hegelian mode, Sellars writes: "the very fact that I use the analogy of stereoscopic vision implies that as I see it the manifest image is not overwhelmed in the synthesis" (SPR: 9). Significantly, this sentence from PSIM thematically resembles, at least in spirit (rather than also in letter), Hegel's insistence in the *Difference* essay that dialectical-speculative judgements concerning unity do not involve any kind of commitment to subsumption that eliminates individuality and difference:

> [t]o cancel established oppositions is the sole interest of reason. But this interest does not mean that it is opposed to opposition and limitation in general; for necessary opposition is *one* factor of life, which forms itself by eternally opposing itself, and in the highest liveliness totality is possible only through restoration from the deepest fission. (DFS; Eng. tr.: 91)

Though Sellars's own position – especially his psychological nominalism – is shorn of Hegel's speculative metaphysical mortgages, Sellars could claim that Adorno is guilty of approaching the subject of unity in *exactly* the sort of way rendered unviable by and running counter to the logic of *Aufhebung*: Adorno is attacking crude and crass positivism; but Sellars is not a crude and crass positivist. Therefore, Adorno's salvo misfires. Only UIHUST involves a commitment to a false *epistemic* totality.

Unlike false epistemic totalities, expressive epistemic totalities involve a conception of a unified whole in which heterogeneous (but not inconsistent) epistemic needs and interests are expressed and also fully developed at no cost to the stability of the whole; if anything, the expression and development of heterogeneous (but not inconsistent) epistemic needs and interests is required to avoid epistemic anomie. The consequence of a false epistemic totality, a crystallisation into well-ordered homogeneous complexes under the steering mechanism of UIHUST, is a crisis situation in our epistemic form of life. This is because the subjective and objective conditions for sense-making risk erosion by increasing patterns of discursive hegemonisation and homogenisation.

Just as *governmental* discipline is directed towards homogenising bodies and sexualities, producing "subjected and practised bodies, 'docile' bodies," (Foucault 1997: 138) the *epistemic* disciplinarity of UIHUST is directed towards homogenising vocabulary and inquiry, producing "disciplinary monotony," subjected and practised minds, 'docile' minds (Foucault 1997: 141).[24]

The Adornian critique of Sellars, then, not only mistakes its target, but, worryingly, also gives rise to debilitating anxieties, stemming from its misdirected allegation, that embedded in Sellars's (Hegelian) commitment to unity is a homogenising drive: fusing the normative category of persons to Peirceish neither represents a sort of "blithely decreed disappearances of individualities" (1973: 325), nor a surrender of the MI. As Christias (2019a: 521) notes, "the normative valence of manifest-image concepts, far from completely disappearing from view, is in fact fully preserved (albeit, with a different – i.e., 'scientific-image'–'contentual' aspect)".

Far from involving epistemic serfdom, the Sellarsian *Aufhebung* of the tension between the MI and SI points to a clear commitment to a more polychromatic, republican pluralism, rather than a monochromatic, imperialist monism. For, "[t]he normative core of the manifest image (the individual and community intentions of persons) is preserved yet completely purged of its ontological-explanatory content, which is now fully accounted for in scientific-image terms" (Christias 2019b: 463-464):

> [w]hen I talk about the in principle replaceability of the manifest image by the scientific image, I do so with respect to the *content* of the world, its material and not with respect to those forms which concern the normative, the obligatory, the correct, the incorrect, the valuable. (WSNDL: 169)

The *Aufhebung* in no way threatens to swamp persons and the language of community and individual intentions, not least because Sellars's functional classification distinction between normative discourse and descriptive-explanatory discourse anchors his Kantian commitment to a 'no-competition,' 'never-the-twain-shall-meet' view of the SI and the conceptual framework of

[24] I think it is important to note here that worries about scientism ought not to use science as a scapegoat for the pathological features of capitalism. Once one sees that pragmatic realism in philosophy of science does not entail – and in fact, strictly speaking, undermines – UIHUST, 'scientism' just becomes a chimera. Given this, the following pertinent question arises: 'why, from a diagnostic perspective, does scientism still persist?' Scientism is, therefore, peculiar, because it persists *despite* resting on implausible grounds, since "the omnipresent neo-Pythagoreanism of contemporary science is surely not adequately justified by its empirical successes" (Dupré 1995: 224). I think a particularly compelling answer to this question involves explaining scientism's persistence in terms of *scientism*'s status – not science's status – as the theoretical concomitant of the kind of social pathologies caused by the ideological exercise of formal reason in capitalist modes of production.

persons. The SI and the conceptual framework of persons have different discursive functions and operate in different discursive domains. MI-discourse is individuated by prescriptive, expressive language that is logically irreducible to descriptive-explanatory vocabulary, even though at the level of ontology, the MI is causally reducible to the descriptive-explanatory categories of natural science. Therefore, to quote deVries (2016a: 58), "[g]iven the structured holism of Sellarsian semantics, the right way to think of the relation between the manifest image and the scientific image is as a matter of mutual accommodation, not one-way dominance" *pace* the Adornian critique.

4.

4.1. The Adornian strikes back: Sellars's Ramsifying slip?

As things currently stand, Adorno *appears* to view Sellars's naturalism through the prism of bad faith: at worst, Sellars is a perfidious pluralist who surreptitiously smuggles scientism through customs;[25] at best, Sellars is naïve, non-wilfully ignorant of the ways in which he reproduces ideology in Western metaphysics and philosophy of mind. Suffice to say that the Sellarsian would view the Adornian critique not necessarily through the prism of bad faith, but at least with a qualified incredulity: the Adornian fails to acknowledge the functional classification distinction and its *Aufhebung* role, and the Adornian conflates Sellarsian naturalism with exactly the sort of naturalism Sellars rejects. However, I think this apparent stalemate is breakable. Specifically, I think the stalemate can be broken in the Adornian's favour, thereby putting significant pressure on the idea that the synoptic vision is an expressive epistemic totality.

I previously claimed that Sellars's functional classification distinction between normative discourse and descriptive-explanatory discourse anchors his Kantian commitment to a 'no-competition,' 'never-the-twain-shall-meet' view of the SI and the conceptual framework of persons. The SI and the conceptual framework of persons have different discursive functions and operate in different discursive domains. MI-discourse is individuated by prescriptive, expressive language that is logically irreducible to descriptive-explanatory vocabulary, even though at the level of ontology, the MI is causally reducible to the descriptive-explanatory categories of natural science. Sellars, as I have argued, maintains that causal reducibility is categorically distinguished from logical

[25] As Adorno puts this in terms of a hermeneutical principle: "what these works say, is not what their words say" (2013: 184).

reducibility, insofar as causal reducibility is solely concerned with descriptive-explanatory power, which is functionally orthogonal to logical irreducibility. The primacy of the SI, therefore, consists in how the SI, rather than the MI, instructs *only* on ontological matters. Understood in this manner, so the story goes, the MI will 'wither away' *only* on the ontological side of sense-making. There is a significant amount riding on just how strict the functional classification distinction is for Sellars, not least because if he Ramsifies across all functional classifications, and moves away from the Kantian functional classification strictures, he runs the real risk of vindicating the Adornian critique.

In PSIM, Sellars was committed to the claim that the SI has primacy over the MI, with respect to instructing the MI on ontological matters. However, to avoid the total colonisation of the MI by scientistic varieties of naturalism, one acts as a conceptual border-patroller, erecting a protective hermeneutic barrier, whose structural integrity one maintains as best as one reasonably can. The functional classification distinction played this very specific epistemic(-political) role, preserving the conceptual autonomy of the MI, thereby keeping the barbarians at the gate (so to speak).

As far as I can tell, there is no textual evidence of Sellars Ramsifying in PSIM or in EPM. *Science and Metaphysics* (S&M) might seem a different story. This is because, at one crucial point, Sellars appears to Ramsify across all functional classifications, and by doing so, moves away from the Kantian functional classification strictures:

> …the Scientific Realist need only argue that a correct account of concepts and concept formation is compatible with the idea that the "language entry" role could be played by statements in the language of physical theory, i.e. that in principle this language could *replace* the common-sense framework in *all* its roles, with the result that the idea that scientific theory enables a more adequate picturing of the world could be taken at its face value. (S&M: V.90)

In writing "that in principle [the language of physical theory] could *replace* the common-sense framework in *all* its roles," Sellars *seems* to Ramsify across all functional classifications. He seems no longer committed to the idea that, with respect to the MI's *expressive* function, the MI is logically irreducible to the descriptive-explanatory categories of postulational science. To use a phrase from Davidson (2001: 207), it now seems that the conceptual framework of persons no longer, at the logical level, resists "capture in the nomological net of physical theory". Sellars's *scientia mensura* doctrine no longer seems to hold for *only* the domain of description and explanation. The Adornian has good reason to feel vindicated now, as this Ramsifying 'slip' appears to confirm their long-standing

suspicion that Sellars's 'eye on the whole' and the synoptic vision reproduce the imperialist modes of thinking embedded in scientific naturalism. The 'eye on the whole' is now revealed as discursively oppressive. The synoptic vision is now revealed as a false epistemic totality. As Jürgen Habermas writes,

> [w]hen stripped of their ideological veils, the imperatives of autonomous subsystems make their way into the lifeworld from the outside – like colonial masters coming into a tribal society – and force a process of assimilation upon it. (1987b: II, 355)

Buoyed by the spectacular and rapid expansions in modern scientific knowledge, scientific naturalism moves from *explanatory* superiority to "regulatory hegemony" (Butler 2009: 5), so much so that inquiries paradigmatically defined by the operation of formal rationality began to *epistemically manage* the MI with a view to seeing it wither away in terms of its ontological *content* and its normative *form*. The synoptic vision, given Sellars's apparent Ramsifying slip in S&M that *all* roles of MI-talk can in principle be replaced by SI-talk, involves the domination of MI discursive forms of life. The conceptual framework of persons and the language of individual and community intentions are, in fact, no longer conceptually autonomous, not so much because they are embedded within the impersonal framework of Peirceish, but because Peirceish ultimately forces the collapse of Sellars's functional classification distinction in favour of descriptive-explanatory categories: central folk psychological concepts no longer remain logically intact in the ideal SI. "The dominance of physics in Sellars's picture, the need to telescope the special sciences into the microphysical sciences, is here disabling" (Levine 2019: 266).

However, it would be premature for the Adornian to proclaim victory. For, Sellars's point is that the language of physical theory replaces the language of the "proper sensibles" (S&M: V.16), in *language-entry* moves. Such a claim does *not* undermine Sellars's commitment to the essentially normative character of material inferences once we are in the space of reasons, because "completely determinate "basic" perceptual this-suches" (S&M: V.16) do not have an epistemic function to begin with *once we give up the Myth of the Given*:[26] proper sensibles do not play an epistemic role in observation reports, following Davidson, and that in turn allows the scientific realist to claim that at least in

[26] The Myth of the Given can be explicated in the following manner: it refers to the traditional empiricist claim that perceptual judgments are epistemically justified by non-conceptual sense contents. At the *base* of our perceptual experience, there are things which do not have propositional content that immediately provide us with epistemic relations, particularly relations of justification. But, perceptual judgments, for Sellars, can only be justified to the extent that they have epistemic relations with cognitive states, things with propositional content. Cf. Sobstyl 2004: 133; O'Shea 2016: 2; Kukla 2006: 85-86.

principle the proper sensibles *could* be replaced by terms from some suitably enlarged language of physical theory. One would talk about light frequencies instead of 'colour;' audio frequencies instead of 'sound;' atoms arranged table-wise instead of 'tables.' *Unless it is essential to the framework of persons that persons see colours, hear sounds, and experience tables*, there is no problem for Sellars here.

In what follows, I argue that rather than resolve the clash between the MI and the SI by joining the 'lifeworldy' conceptual framework of persons to the SI for the purpose of enriching and completing the SI, what Sellars ought to have done is adopt a negative dialectical 'resolution' of the clash between the images. This strategy invites one to dismantle the Placement Problem[27] *qua* "a logic of disintegration" (Adorno 2008: 6). I take Sellars to have curiously hinted at – but not follow through – this Adornian intellectual orientation in EPM.

4.2. Disintegration: "an *arché* beyond discourse"

What can Sellars do now (at least from a left-wing Sellarsian perspective)? There are several paths open to him. I wish, though, to focus on one path which, were Sellars to take it, would not cause any kind of reputational embarrassment for him:[28] Sellars can perform a conceptual about-turn, and he can disavow his claim in *Science and Metaphysics* that *all* roles of MI-talk can in principle be replaced by SI-talk; he can return to his original position in PSIM, namely that the way to resolve the clash between the MI and the SI is to *integrate* the conceptual framework of persons with the 'doggedly naturalistic' language of postulational science. By joining the language of individual and community intentions to Peirceish, such a discursive activity enriches and completes the SI. To repeat Christias's characterisation of the Sellarsian *Aufhebung*, the idea here is 'a smooth incorporation of normativity within the scientific image.'

Should this be the path Sellars takes, I think the Adornian can up the ante in the high-stakes game here precisely because Sellars seems so Hegelian: why should we even seek to *join* the conceptual framework of persons to Peir-

[27] The Placement Problem can be formulated in this manner: (1) All reality is ultimately natural reality. (2) Whatever one wishes to admit into natural reality must be placed in natural reality. (3) Modality, meaning, norms, consciousness, self-consciousness, and intentionality, and so on do not seem admissible into natural reality. (4) Therefore, if they are to be placed in nature, they must be forced into a category that does not seem appropriate for their specific characters; and if they cannot be placed in nature, then they must be either dismissed as non-genuine phenomena, or at best regarded as parasitic second-rate phenomena.

[28] If anything, it might exemplify a Putnam-esque self-critical turn, and re-confirm Sellars's pragmatist disposition.

ceish? Furthermore, why think the dialectical aspiration ought to be *smooth* incorporation? And, even more basically, why think the *telos* of our cognitive endeavours, why contend that the goal of inquiry is to *end the internal tension* between the two images?

My Adornian line of thought principally involves construing the discourse of joining the conceptual framework of persons to the SI for the purpose of enriching and completing it as, what I would term, 'axiologically corrosive.' For, to construe the MI's value ultimately in terms of how it benefits the SI, to the extent that affixing the language of individual and community intentions to Peirceish enriches and completes the SI, seems invariably committed to the idea that the model of rationality in Peirceish – "our more straitlaced" cousin (N&O: 6) – occupies the centre of our sense-making web. Specifically, the idea that the model of rationality in Peirceish occupies the centre of our sense-making web means that the conceptual framework of persons has to *serve* Peirceish (precisely by the unidirectionality of enriching and completing the SI). This reinforces the ideological-regulatory structural features of the Placement Problem, as well as revealing how such disciplinary features bleed into those long-standing liberal naturalist responses to it.[29] Perhaps more polemically, the axiological corrosiveness of the joining-discourse involves, to quote Max Weber, the 'mechanised petrification' of our sense-making itself – the extirpation of person-practices.

The Placement Problem aims to level out the heterogeneous dimensions of the MI, by framing the legitimacy of scientifically recalcitrant phenomena in terms of whether they can be placed/located in the world described by the natural sciences: anything that resists placeability/locatability is labelled 'odd.' By being *visibly marked*, 'odd' phenomena become 'queer' phenomena, which then become 'problematic' and 'punishable' phenomena. There is compelling reason to think that nomothetic structure of placeability/locatability operate *juridically*. Paraphrasing Judith Butler on juridical operationality, "the subjects regulated by such structures are, by virtue of being subjected to them, formed, defined, and reproduced in accordance with the requirements of those structures" (Butler 1999: 4).

Regulatory-juridical logical structures and their corresponding axiological commitments are operative in the allegedly capacious discourse of 'accommo-

[29] The broad tradition of liberal naturalism *as an intellectual orientation for coordinating non-eliminativist, non-reductionist discourse about normative kinds* often claims to have significant advantages over its more conservative (or scientific naturalist) cousin. Importantly, liberal naturalists *explicitly* maintain their naturalist credentials, but do so in such a way that aims to make a clear demarcation between them, supernaturalists, and scientific naturalists. Viz. Giladi 2019b, De Caro and Macarthur 2010: 9.

dation': expressions such as 'finding a place for mind in the world described by the natural sciences' and 'making elbow room for intentionality' both presuppose that one ought to accept from the very outset the vocabulary and general *Weltanschauung* of the natural sciences, and then find some meaningful and coherent way of fitting/affixing intentionality into that nomothetic picture. This is axiologically corrosive of our sense-making itself, because the language of individual and community intentions and the conceptual framework of persons, where it is essential to *this* framework that persons see *colours*, hear *sounds*, experience *tables*, have principally intrinsic, not relational, value: they are expressive of the lifeworldy practices we engage in to make sense of *Erlebnis* as an intentional, communicative, socially invested agent (cf. Bernstein 1966: 15). In this way, the task at hand is not to find ways of accommodating intentionality. Rather, the task is to combat the circulation of *epistemic power.* This way of thinking here, to quote Michel Foucault,

> should be seen as a kind of attempt to emancipate historical knowledges from that subjection, to render them, that is, capable of opposition and of struggle against the coercion of a theoretical, unitary, formal and scientific discourse. It is based on a reactivation of local knowledges – of minor knowledges, as [Gilles] Deleuze might call them – in opposition to the scientific hierarchisation of knowledges and the effects intrinsic to their power. (1980: 85)

Interestingly, while not laced in critical theoretic terms, Sellars's own normative pragmatism about knowledge, where epistemic kinds are made sense of not *qua* conceptual analysis, but *qua* a recognisable standing in the space of reasons, is *allied* to Foucault's point: in not giving an empirical description of epistemic kinds, Sellars offers a revisionary epistemology, to the extent that normative pragmatism about knowledge is a 'reactivation of local knowledges.'

These local knowledges include 'knowing one's way around (the space of reasons)' and 'rehearsing intentions,' which comprise those practices involved in the intersubjectively constituted 'ought-to-do's' and 'ought-to-be's.' They are 'local' in the sense of localised in the full-range of practices only performable by *persons*. As Sellars makes it clear, the pragmatically salient features of sensitivity to a normative *community*, particularly the sensitivity to rules of criticism, is something only predicable of *persons.* The reactivation of local knowledges, such as deontic scorekeeping and its concomitant processes of agonistic constitution of 'ought-to-be's' and 'ought-to-do's,' reveals the normative lustre of the space of reasons and the conceptual framework of persons: as persons, we are not just rule-governed; we are also normatively self-constituting as *persons*. In other words, we construe our practical identity/our practical relation-to-self

as agents whose speech acts and actions in general are structured by reasons in accordance with rules. By having our speech acts and actions in general structured by reasons in accordance with rules, all of which are moulded in the crucible of the space of reasons, *we constitute ourselves as persons.*

Applying conceptual analysis to this kind of discourse is a category error, because conceptual analysis is "too buttoned-up and white-chokered and clean-shaven a thing" (EIRE: 146) to adequately make sense of, for example, the norms governing knowledge-attribution and the agonistic constitution of rules of criticism. Sellars, as a pragmatist, is decidedly uninterested in finding any features/states/properties serving as *formal* conditions of knowledge. The norms governing knowledge-attribution, the agonistic constitution of rules of criticism, etc. all involve deliberative discourse in order to be authoritative, legitimate, and valid for those engaging in such discourse. By conceptualising knowledge in terms of a recognisable standing in the logical space of reasons, the "network of discursive holdings" (Kukla and Lance 2009: 192), Sellars's pragmatism is decidedly interested in the *informal*, *flexible*, and *humanistic* norm-constituting practices of language-using agents.

I think that for all of Sellars's emphasis on the rule-governed features of human language and action, the informal, flexible, and humanistic norm-constituting practices of persons, crucially, involve *opposition* and *struggle*, so much so that, as Joseph Margolis writes,

> [y]ou must bear in mind that "*to place an item* in a normative space" (as Sellars has it) *is* to place *it* (consulting doxastic or cognitive attributions that characteristically trigger nonmonotonic complications) in a decidedly uncertain – possibly unmanageable – inferential space. (2016: 20)

A 'decidedly uncertain – possibly unmanageable –' inferential space is decidedly uncertain and possibly unmanageable – precisely because *the space of reasons is an arena invariably comprising opposition and struggle, contestation and challenge, disruption and disturbance.* Significantly, for the Adornian, opposition and struggle, contestation and challenge, disruption and disturbance are the effects of the ineliminable presence of *non-identity* in the conceptual framework of persons: most importantly, this category *eo ipso* puts the brakes on the Sellarsian idea of "an ever-expanding range of homeostatic equilibrium" (Christias 2019b: 465). If anything, this very notion of an ever-expanding range of homeostatic equilibrium or allostatic regulation – the smooth incorporation of normativity within the SI – is *precisely* what concerns the Adornian, and makes Sellars vulnerable to the Adornian critique of Hegel, given how deeply Hegelian Sellars himself was:

contradiction cannot be brought under any unity without manipulation, without the insertion of some wretched cover concepts that will make the crucial differences vanish. (1973: 152)

The ineliminable presence of non-identity means that the dialectical aspiration never ought to be an incorporation of the conceptual framework of persons to Peirceish (let alone a 'smooth' one); the *telos* of our cognitive endeavours never ought to end the internal tension between the two images. The ineliminable presence of non-identity means that the dialectical aspiration always ought to *bathe in the clash* between the MI and the SI; the *telos* of our cognitive endeavours always ought to *emphasise internal tension* between the two images.

Construed in the manner I have articulated, I think it is plausible to argue that just as Adorno himself wished to *rescue* Hegel (viz. Adorno 1993 [1963]: 83), the Adornian wishes to *rescue* Sellars, so that the Sellarsian *Aufhebung* can overcome its fear of non-identical thinking. In this way, much of the following by Hammer, which focuses on, but risks underplaying the *force* of, Adorno's deployment of negative dialectics in social theory against ideological integration, carries over to my Adornian worry about smoothly incorporating the conceptual framework of persons within Peirceish:

[s]ince the *modern* social systems within which we find ourselves so strongly ideologically (and hence 'misleadingly') encourage us to believe that a successful and meaningful integration [...] has indeed taken place, the reconciliatory step towards *Aufhebung* should be resisted in favour of a focus on the *distance* between notional constraint and our ways of knowing and relating to the world (Hammer 2020: 40).

The function of negative dialectics is not to offer *resistance* to the totalising dispositions of modern social organisation and scientific hierarchisation of knowledges. Rather, the function of negative dialectics is to *reverse* the direction of discursive power and dismantle the Placement Problem[30] through the

[30] In recent years, the Placement Problem has been critiqued by philosophers of either (i) a Hegelian inclination, who try to dissolve the Placement Problem by articulating how it rests on the non-dialectical framework of *Verstand* (as opposed to the dialectical framework of *Vernunft*) – see Giladi 2014; 2019a; or (ii) a (neo-)Kantian inclination, who try to dissolve the Placement Problem by showing how it is based on presuppositions that fail to underpin different forms of experience and (therefore) different ways of knowing – see D'Oro 2018; 2019 and Papazoglou 2019; or (iii) a Husserlian inclination, who try to dissolve the Placement Problem using the perspective of transcendental phenomenology – see Moran 2008; 2012; 2013 and Hanna 2014; or (iv) a Wittgensteinian inclination, who try to dissolve the Placement Problem by showing how it distorts the relationship between grammar and experience, conflating saying and showing – see Beale and Kidd 2017; or (vi) a broadly pluralist realist inclination, who try to dissolve the Placement Problem by relaxing the notion of nature in such a way

exercise of a logic of disintegration. There are, I think, at least two ways the Sellarsian can respond here.

In the first instance, they might reply that the Adornian, ironically, wishes to maintain hierarchisation and domination by now forcing the SI to conform to the MI. This merely reproduces the Placement Problem in a reverse form: the conceptual framework of the natural sciences has to be placed/located in the messy normative space of reasons on pain of humanistic (rather than naturalistic) Ramsification. The Adornian position, then, is, at best, hypocritical; at worst, absurd.

In the second instance, which I think is the much better dialectical path to take here, the Sellarsian may point to the concluding sentence of EPM:

> Or does the reader not recognise Jones as Man himself in the middle of his journey from the grunts and groans of the cave to the subtle and polydimensional discourse of the drawing room, the laboratory, and the study, the language of Henry and William James, of Einstein and of the philosophers who, in their efforts to break out of discourse to an *archê* beyond discourse, have provided the most curious dimension of all. (SPR: 196; emphasis added)

The 'efforts to break out of discourse to an *archê* beyond discourse' curiously hints at Adorno's negative dialectical orientation. This is because the logic of Sellars's expression here bears noticeable resemblance to Adorno's logic of disintegration, whose aim is "[t]o use the strength of the subject to break through the fallacy of constitutive subjectivity" (1973: 10). To achieve success in philosophy would be, then, to 'know one's way around' *with respect to internal tension*, rather than with respect to welding into one unified, coherent image. Putting Sellars and Adorno into conversation with one another provides this most curious dimension, namely that our discursive forms of life require multiple images, multiple pictures, which are in conflict with one another, because conflict, rather than a transcending *Aufhebung*, is emblematic of cognitive life itself (cf. Christias 2018b: 128).

However, if what I have claimed here is correct, then one may legitimately wonder what is holding Sellars back from following through the Adornian intellectual disposition here. To put the point more bluntly, why does the concluding sentence of EPM 'hint at' rather than 'explicitly articulate' a negative

that removes the spectre of reduction or elimination – see Baker 2013; 2017; McDowell 1994; Putnam 1990; 1994; 1995; 2002; 2004; 2012; 2015; De Caro 2015; 2019; or (vii) a Rortian neopragmatist inclination, who try to dissolve the Placement Problem by revealing how it is produced by representationalist, rather than expressivist, grammar, namely the idea that semantics and our discursive vocabulary involve a mirroring word-object relationship – see Rorty 2010; Price 2004; Macarthur and Price 2007; Macarthur 2008.

dialectical orientation? I think a particularly helpful way to answer this question involves re-emphasising how Hegelian Sellars is, and therefore creatively (in a non-Rortian way) reanimating (i) some of Adorno's critique of Hegel in *Hegel: Three Studies* (1993 [1963]) and some of Habermas's critique of Hegel in the *Philosophical Discourse of Modernity* (1987a).

Adorno and Habermas respectively write that

> [a]s though the dialectic had become frightened of itself, in the *Philosophy of Right* Hegel broke off such thoughts by abruptly absolutising one category – the state. This is due to the fact that while his experience did indeed ascertain the limits of bourgeois society, limits contained in its own tendencies, as a bourgeois idealist he stopped at that boundary because he saw no real historical force on the other side of it. He could not resolve the contradiction between his dialectic and his experience: it was this alone that forced Hegel the critic to maintain the affirmative. (Adorno 1993 [1963]: 80)

The point of the intuitions from the days of his youth that Hegel wanted to conceptualise was that in the modern world emancipation became transformed into unfreedom because the unshackling power of reflection had become autonomous and now achieved unification only through the violence of a subjugating subjectivity (1987a: 32-33).

As we have seen, in Hegel's youthful writings the option of explicating the ethical totality as a communicative reason embodied in intersubjective life-contexts was still open. Along this line, a democratic self-organisation of society could have taken the place of the monarchical apparatus of the state. By way of contrast, the logic of a subject conceiving itself makes the institutionalism of a strong state necessary … Hegel had hardly conceptualised the diremption of modernity before the unrest and movement of modernity was ready to explode this concept. The reason for this is that he could carry out his critique of subjectivity only within the framework of the philosophy of the subject (1987a: 40-41).

For Adorno, there is a clear distinction between Hegel and Hegelianism, where 'Hegelianism' refers not so much to Hegel's actual philosophical commitments and arguments, but principally to a very specific constellation of conceptual frameworks, methodologies, and discursive resources. True to the spirit of the Young Hegelians, Adorno sees Hegelianism as comprising the necessary methodological principles as well as the materialist discursive tools for sustained and progressive *social critique*. More directly put, Hegelianism unshackles Hegel's dialectic from ideology and *Begriffsbildung*, so that Hegel's dialectic can overcome its long-standing fear of non-identical thinking. And Habermas lambasts the mature Hegel on the grounds that he "did not pursue

any further the traces of communicative reason that are clearly to be found in his early writings" (1987a: 31). Instead, according to Habermas, using Dieter Henrich's expression, Hegel articulated an aloof "emphatic institutionalism" (1987a: 41) in the *Philosophy of Right*, leaving the critical concepts of intersubjectivity and communicative action underdeveloped and their emancipatory potential in stasis.

How exactly does this bear on Sellars, though? The critique of Hegel partly revolves around the contention that Hegel is intoxicated by manageability, by the monistic vision of a dialectically integrated Absolute, to the point where the dialectical process mythopoetically presses the *need* for a unified and coherent whole. However, as Adorno remarks,

> but hidden in it is also the true moment of ideology, the pledge that there should be no contradiction, no antagonism. (1973: 149)

In this way, the charm of the monistic vision's dialectically integrated Absolute is unmasked as pathologically bewitching. Similarly, the critique of Sellars partly revolves around the contention that Sellars is smitten by manageability, by the monistic vision of a dialectically integrated naturalistic image of the world, to the point where the dialectical process here mythopoetically presses the *need* for a unified and coherent naturalistic vision.

There is something almost irresistible to naturalism. Crucially, what makes naturalism so appealing is that its charming qualities deeply resonate with our psychological architecture and cognitive make-up: as human beings, we are sense-making creatures. We inquire to render the world around us rationally intelligible. From an anthropological perspective, then, naturalism's "charm" (see Stroud 1996) consists in appealing to our basic cognitive drive to render reality discursively manageable. Under such a conception of naturalism, the idea of bringing into question such an orientation of thinking, one which taps into our need as a species to rationally make sense of things, seemingly countenances blocking the way of inquiry. This would be anathematic to the very function of philosophical reflection, leaving *reality* not only discursively inaccessible, but also leaving us radically alienated from our *own* nature. Furthermore, naturalism is not just charming at the primitive anthropologic-psychological level; naturalism is also appealing because of just how successful and emancipatory the natural sciences have undeniably been. Questioning naturalism, then, would be tantamount to disputing the remarkable epistemic successes of physicists, chemists, and biologists. As C.S. Peirce famously wrote, "[a] man must be downright crazy to deny that science has made many true discoveries" (Pierce 1992: II, 217).

However, the charm of the monistic vision's dialectically integrated single, unified, coherent naturalistic image is unmasked as pathologically bewitching: Peirceish – "our more straitlaced" cousin – occupies the centre here, and its naturalistic categories are the products of the domination, not revelation, of nature. The enriched and completed naturalism of the ideal SI is not as innocent as Sellars makes it out to be.

Sellars is held back from following through his Adornian intellectual disposition at the conclusion of EPM by valorising manageability, by his advocacy of "[t]he primacy of totality over phenomenality" (1973: 303). In disfavouring paradox, seeking fairly neat solutions, and glorifying smooth integration, one is not discursively protected against the damage to forms of cognitive life. Disfavouring paradox, seeking fairly neat solutions, and glorifying smooth integration comprises "the defamation of alternative modes of thought which contradict the established universe of discourse" (Marcuse 2002: 178).

That Sellars holds himself back reveals some type of anxiety, perhaps even a fear of "radically new forms of sense-making" (Moore 2012: 192), a fear of creativity, a fear of the unmanageable, a reactionary disposition to genuinely challenging and even overcoming the discursive *status quo*.[31] The 'efforts to break out of discourse to an *arché* beyond discourse' "would be the concept of an open dialectic – in contrast to the closed dialectic of idealism" (ID: 21). Taken this way, when Christias (2018a: 1317) asks if it is "plausible to suggest that our conception of what we really are is bound to be *necessarily* fragmented?", from the Adornian perspective I have advocated, the answer is not just that it is plausible, but that it is *true*. As Jerry Fodor (1997: 162) writes, "[y]ou may find that perplexing; you certainly aren't obliged to like it. But I do think we had all better learn to live with it". If the Sellarsian stereoscopic vision involves bottling Hegelian wine in naturalistic bottles, the idealist longing for totality results in corking the wine.

Paul Giladi
Manchester Metropolitan University
p.giladi@mmu.ac.uk

[31] I am very grateful to Dionysis Christias, Carl Sachs, and Steven Levine for their patient, detailed, and constructive criticisms of earlier drafts of this paper. Their warm dialectical engagement with my arguments has significantly improved my own grasp of the complexities of Sellars's position, and enabled me to develop a more articulated and nuanced evaluative perspective.

References

Abbreviations for works by Wilfrid Sellars

CDCM = 1958, "Counterfactuals, Dispositions, and the Causal Modalities", in H. Feigl, M. Scriven, and G. Maxwell, eds. *Concepts, Theories, and the Mind-Body Problem*, University of Minnesota Press, Minneapolis (MN): 225-308.

EPM = 1997, *Empiricism and the Philosophy of Mind*, Harvard University Press, Cambridge (MA).

N&O = 1980b, *Naturalism and Ontology*: *The John Dewey Lectures for 1973-74*, Ridgeview Publishing Co, Reseda (CA).

PP = 1977, *Philosophical Perspectives: Metaphysics and Epistemology*, Ridgeview Publishing Co, Atascadero (CA).

PPPW = 1980a, *Pure Pragmatics and Possible Worlds: The Early Essays of Wilfrid Sellars*, Ridgeview Publishing Co, Atascadero (CA).

S&M = , 1968, *Science and Metaphysics: Variations on Kantian Themes*, Routledge, London.

SPR = *Science, Perception and Reality*, Routledge, London & New York.

WSNDL = 2015, *Wilfrid Sellars: Notre Dame Lectures 1969-1986*, ed. by P.V. Amaral, Ridgeview Publishing Co, Atascadero (CA).

Other works

Adorno, Theodor Wiesengrund, 1973, *Negative Dialectics*, Eng. tr. by E. B. Ashton, Continuum, New York.

Adorno, Theodor Wiesengrund, 1974, *Minima Moralia: Reflections from Damaged Life*, Eng. tr. by E. Jephcott, Verso, London.

Adorno, Theodor Wiesengrund, 1993, *Hegel: Three Studies*, Eng. tr. by S.W. Nicholsen, MIT Press, Cambridge (MA).

Adorno, Theodor Wiesengrund, 2008, *Lectures on Negative Dialectics: Fragments of a lecture course 1965/1966*, ed. by R. Tiedemann, Eng. tr. by R. Livingstone, Polity Press, Cambridge (MA).

Adorno, Theodor Wiesengrund, 2013, *Aesthetic Theory*, ed. by G. Adorno and R. Tiedemann, Eng. tr. by R. Hullot-Kentor, Bloomsbury, London.

Adorno, Theodor Wiesengrund, 2017, *Introduction to Dialectics*, ed. by C. Ziermann, Eng. tr. by N. Walker, Polity Press, Cambridge (MA).

Baker, Lynne Rudder, 2013, *Naturalism and the First-Person Perspective*, Oxford University Press, Oxford.

Baker, Lynne Rudder, 2017, "Naturalism and the Idea of Nature", in *Philosophy*, 92: 333-349.

Beale, Jonathan, and Kidd, Ian James, eds. 2017, *Wittgenstein and Scientism*, Routledge, New York.

Bernstein, Richard J., 1966, "Sellars's Vision of Man-in-the-Universe I", in *Review of Metaphysics*, 20: 113-143.

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

Butler, Judith, 1999, *Gender Trouble: Feminism and the Subversion of Identity*, 2nd Edition, Routledge, London & New York.

Carnap, Rudolf [1958], 1975, "Observational Language and Theoretical Language", in Jaako Hintikka,ed. *Rudolf Carnap, Logical Empiricist: Materials and Perspectives*, Springer, Dordrecht: 75-85.

Cartwright, Nancy, Cat, Jordi, Fleck, Lola, and Uebel, Thomas E., 1996, *Otto Neurath: Philosophy Between Science and Politics*, Cambridge University Press, Cambridge (MA).

Cavarero, Adriana, 2016, *Inclinations: A Critique of Rectitude*. Eng. tr. by A. Minervini and A. Sitze, Stanford University Press, Stanford (CA).

Christias, Dionysis, 2016a, "Can 'Ready-to-Hand' Normativity be Reconciled with the Scientific Image?", in *Philosophia*, 44: 447-467.

Christias, Dionysis, 2016b, "Can Sellars's argument for scientific realism be used against his own *scientia mensura* principle?", in *Synthese*, 193: 2837-2863.

Christias, Dionysis, 2018a, "On the proper construal of the manifest-scientific image distinction: Brandom *contra* Sellars", in *Synthese*, 195: 1295-1320.

Christias, Dionysis, 2018b, "Reconciling Scientific Naturalism with the Unconditionality of the Moral Point of View: A Sellars-inspired Account", in *Res Philosophica*, 95: 111-149.

Christias, Dionysis, 2019a, "Towards a Reformed Liberal *and* Scientific Naturalism", in *Dialectica*, 73: 507-534.

Christias, Dionysis, 2019b, "The Non-Conceptual Dimension of Social Mediation: Towards a Materialist *Aufhebung* of Hegel", in *International Journal of Philosophical Studies*, 27: 448-473.

Dach, Stephanie, 2018, "Sellars's Two Images as a Philosopher's Tool", in *Metaphilosophy*, 49: 568-588.

Davidson, Donald, 2001, "Mental Events", in D. Davidson,ed.) *Essays on Actions and Events*, Oxford University Press, Oxford.

De Caro, Mario, 2015, "Realism, Common Sense, and Science", in *The Monist*, 98: 197-214.

De Caro, Mario, 2019, "Common-sense and Naturalism", in Giladi 2019b: 184-207.

De Caro, Mario and Macarthur, Douglas, 2010a, *Naturalism and Normativity*, Columbia University Press, New York.

De Caro, Mario and Macarthur, Douglas, 2010b, "Introduction: Science, Naturalism, and the Problem of Normativity", in De Caro and Macarthur 2010a: 1-20.

D'Oro, Giuseppina, 2018a, "The Touch of King Midas: Collingwood on why actions are not events", in *Philosophical Explorations* 21: 1-10.

D'Oro, Giuseppina, 2019, "Between Ontological Hubris and Epistemic Humility: Collingwood, Kant and Transcendental Arguments", in *British Journal of the History of Philosophy*, 27: 336-357.

deVries, Willem A., 2005, *Wilfrid Sellars*, McGill-Queen's University Press, Montreal and Kingston.

deVries, Willem A., 2016a, "Images, Descriptions, and Pictures: Personhood and the Clash", in J.R. O'Shea,ed. *Wilfrid Sellars and His Legacy*, Oxford University Press, Oxford: 47-59.

deVries, Willem A., 2016b, "Just What is the Relation between the Manifest and the Scientific Images?", in *International Journal of Philosophical Studies*, 24: 112-128.

deVries, Willem A., 2017, "Hegelian Spirits in Sellarsian Bottles", *Philosophical Studies*, 174: 1643-1654.

deVries, Willem A., 2019, "Hegel and Sellars on the Unity of Things", in *International Journal of Philosophical Studies*, 27: 363-378.

Finlayson, James G., 2014, "Hegel, Adorno and the Origins of Immanent Criticism", in *British Journal for the History of Philosophy*, 22: 1142-1166.

Fodor, Jerry, 1974, "Special Sciences, or the Disunity of Science as a Working Hypothesis", in *Synthese*, 28: 97-115.

Fodor, Jerry, 1997, "Special Sciences: Still Autonomous After All These Years", in *Noûs*, 31: 149-163.

Foucault, Michel, 1980, *Power/Knowledge*, ed. by C. Gordon, Pantheon, New York.

Foucault, Michel, 1997, *Discipline and Punish: The Birth of the Prison*, Eng. tr. by A. Sheridan, Vintage Books, New York.

Giladi, Paul, 2019a, "The Placement Problem and the Threat of Voyeurism", in Giladi 2019b: 71-97.

Giladi, Paul, 2019b, ed. *Responses to Naturalism: Critical Perspectives from Idealism and Pragmatism*, Routledge, New York.

Giladi, Paul, 2020a, "A Foucauldian Critique of Scientific Naturalism: 'Docile Minds'", in *Critical Horizons*, 21: 264-286.

Habermas, Jürgen, 1971, *Knowledge and Human Interests*, Eng. tr. by J.J. Shapiro, Beacon Press, Boston.

Habermas, Jürgen, 1987a, *The Philosophical Discourse of Modernity: Twelve Lectures*, Eng. tr. by F.G. Lawrence, MIT Press, Cambridge (MA).

Habermas, Jürgen, 1987b, *The Theory of Communicative Action.* Vol. II: *Lifeworld and System*, Eng. tr. by T. McCarthy, Beacon, Boston.

Hammer, Espen, 2020, "The Antinomy of Modernism and Anti-Modernism in Adorno's *Negative Dialectics*", in P. Giladi, ed. *Hegel and the Frankfurt School*, Routledge, New York: 33-52.

Hanna, Robert, 2014, "Husserl's Crisis and Our Crisis", in *International Journal of Philosophical Studies*, 22: 752-770.

Hegel, Georg Wilhelm Friedrich, DFS = "Die Differenz des Fichteschen und Schellingschen Systems der Philosophie", in *Kritisches Journal der Philosophie*; Eng. tr. by H. S. Harris and W. Cerf 1977 *The Difference Between Fichte's and Schelling's System of Philosophy*, State University of New York Press, Albany (NY).

Hegel, Georg Wilhelm Friedrich, EPW = *Enzykopädie der philosophischen Wissenschaften im Grundrisse*; Eng. tr. Part 1 by T. F. Geraets, W. A. Suchting, and H. S. Harris 1991, *The Encyclopaedia Logic*, Hackett, Indianapolisi (IN); Eng. tr. Part 2 by A.V. Miller 1970, *Philosophy of Nature (Part Two of the Encyclopaedia of Philosophical Sciences)*, Oxford University Press, Oxford.

Hicks, Michael, forthcoming, "Wilfrid Sellars and the task of philosophy", in *Synthese*.

Jackson, Franck C., 1998, *From Metaphysics to Ethics: A Defence of Conceptual Analysis* Oxford University Press, Oxford.

James, William, 2003, *Essays in Radical Empiricism*, Dover Publications, London-New York.

Kelly, Erin, 2008, "Doing without Desert", in M. McKenna and P. Russell, eds. *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's "Freedom and Resentment"*, Ashgate, Farnham: 273-296.

Kukla, Rebecca, 2006, "Objectivity and Perspective in Empirical Knowledge", in *Episteme*, 3: 80-95.

Kukla, Rebecca and Lance, Mark, 2009, *'Yo!' and 'Lo!' The Pragmatic Topography of the Space of Reasons*, Harvard University Press, Cambridge (MA).

Levine, Steven, 2019, "Rethinking Sellars's Naturalism", in Giladi 2019b: 250-270.

Lewis, David, 1970, "How to Define Theoretical Terms", *The Journal of Philosophy*, 67: 427-446.

Macarthur, David, 2008, "Pragmatism, Metaphysical Quietism, and the Problem of Normativity", in *Philosophical Topics*, 36: 193-209.

Macarthur, David and Price, Huw, 2007, "Pragmatism, Quasi-realism and the Global Challenge", in C. Misak, ed. *The New Pragmatists*, Oxford University Press, Oxford: 91-121.

Marcuse, Herbert, 2002, *One-Dimensional Man*, Eng. tr. by , Routledge, London/New York.

Margolis, Joseph, 2016, "Persons as Natural Artifacts", in *Pragmatism Today*, 7: 8-22.

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

Mignolo, Walter D., 2011, *The Darker Side of Western Modernity: Global Futures, Decolonial Options*, Duke University Press, Durham (NC).

McDowell, John, 1994, *Mind and World*, Harvard University Press, Cambridge (MA).

Moore, Adrian William. 2012, *The Evolution of Modern Metaphysics: Making Sense of Things*, Cambridge University Press, Cambridge.

Moran, Dermot, 2008, "Husserl's Transcendental Philosophy and the Critique of Naturalism", in *Continental Philosophy Review*, 41: 401-425.

Moran, Dermot, 2012, *Husserl's Crisis of the European Sciences and Transcendental Phenomenology: An Introduction*, Cambridge University Press, Cambridge.

Moran, Dermot, 2013, "'Let's Look at It Objectively': Why Phenomenology Cannot be Naturalised", in *Royal Institute of Philosophy Supplement*, 72: 89-115.

Nagel, Ernest, 1961, *The Structure of Science*, Harcourt, Brace and World, New York.

O'Connor, Brian, 2004, *Adorno's Negative Dialectic: Philosophy and the Possibility of Critical Rationality*, MIT Press, Cambridge (MA) and London.

O'Connor, Brian, 2013. *Adorno*, Routledge, New York.

Oppenheim, Paul and Putnam, Hilary, 1958, "The Unity of Science as a Working Hypothesis", in H. Feigl *et al.*, eds. *Minnesota Studies in the Philosophy of Science.* Vol. 2, Minnesota University Press, Minneapolis: 3-36.

O'Shea, James R., 2007, *Wilfrid Sellars: Naturalism with a Normative Turn*, Polity, Cambridge.

O'Shea, James R., 2009, "On the Structure of Sellars's Naturalism with a Normative Turn", in W. A. deVries,ed. *Empiricism, Perceptual Knowledge, Normativity, and Realism: Essays on Wilfrid Sellars*, Oxford University Press, Oxford: 187-210.

O'Shea, James R., 2016, "What to Take Away from Sellars's Kantian Naturalism", in J.R. O'Shea, ed. *Wilfrid Sellars and His Legacy*, Oxford University Press, Oxford: 130-148.

Papazoglou, Alexis, 2019, "The Idealist Challenge to Naturalism", in Giladi 2019b: 98-119.

Pierce, Charles S., 1992, *The Essential Peirce: Selected Philosophical Writings*, ed. by. Nathan Houser and Christian Kloesel, 2 vols., Indiana University Press, Bloomington (IN).

Price, Huw, 2004, "Naturalism without Representationalism", in De Caro and Macarthur 2010a: 71-88.

Putnam, Hilary, 1990, *Realism with a Human Face*, Harvard University Press, Cambridge (MA).

Putnam, Hilary, 1994, *Words and Life*, Harvard University Press, Cambridge (MA).

Putnam, Hilary, 1995, *Pragmatism: An Open Question*, Blackwell, Oxford.

Putnam, Hilary, 2002, *The Collapse of the Fact/Value Dichotomy and Other Essays*, Harvard University Press, Cambridge (MA).

Putnam, Hilary, 2004, "The Content and Appeal of 'Naturalism'", in De Caro, Mario and Macarthur, Douglas, eds. *Naturalism in Question*, Harvard University Press, Cambridge (MA): 59-70.

Putnam, Hilary, 2012, *Philosophy in an Age of Science: Physics, Mathematics, and Scepticism*, ed. by. M. De Caro and D. Macarthur, Harvard University Press, Cambridge (MA).

Putnam, Hilary, 2015, "Naturalism, Realism, and Normativity", in *Journal of the American Philosophical Association*, 1: 312-328.

Quine, Willard V., 1948, "On What There Is", in *The Review of Metaphysics* 2: 21-38.

Ramsey, Franck P., 1931, *The Foundations of Mathematics and Other Logical Essays*, ed. by R.B. Braithwaite, Kegan Paul, London.

Rescher, Nicholas, 2017, "Philosophy as Rational Systematisation", in G. D'Oro and S. Overgaard, eds. *The Cambridge Companion to Philosophical Methodology*, Cambridge University Press, Cambridge: 32-43.

Rorty, Richard, 1997, "Introduction", in Sellars 1997: 1-12.

Rorty, Richard, 2010, "Naturalism and Quietism", in De Caro and Macarthur 2010a: 55-68.

Rosenberg, Alexander, 2014, "Disenchanted Naturalism", in B. Bashour and H.D. Muller, eds. *Contemporary Philosophical Naturalism and Its Implications*, Routledge, New York/London: 17-36.

Rosenberg, Jay F., 2007, *Wilfrid Sellars: Fusing the Images*, Oxford University Press, Oxford.

Rosenberg, Jay F., 2009, "On Sellars's Two Images of the World", in W.A. deVries, ed. *Empiricism, Perceptual Knowledge, Normativity, and Realism: Essays on Wilfrid Sellars*, Oxford University Press, Oxford.

Rosenthal, David, 2016, "Quality Spaces, Relocation, and Grain", in J.R. O'Shea, ed. *Wilfrid Sellars and His Legacy*, Oxford University Press, Oxford.

Sobstyl, Edrie, 2004, "Re-radicalising Nelson's Feminist Empiricism", in *Hypatia*, 19: 119-141.

Stern, Robert, 2009, *Hegelian Metaphysics*, Oxford University Press, Oxford.

Strawson, Peter F., 2008, "Freedom and Resentment", in M. McKenna and P. Russell, eds. *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's "Freedom and Resentment"*, Ashgate, Farnham: 19-36.

Wedgwood, Ralph, 2007, *The Nature of Normativity*, Oxford University Press, Oxford.

Wolf, Susan, 2008, "The Importance of Free Will", in M. McKenna and P. Russell, eds. *Free Will and Reactive Attitudes: Perspectives on P. F. Strawson's "Freedom and Resentment"*, Ashgate, Farnham: 69-83.

Young, Iris Marion, 1990, *Justice and the Politics of Difference*, Princeton University Press, Princeton (NJ).

Past Present

The significance of *Quasizerlegung* for Carnap's *Aufbau* and scientific philosophy in general

Caterina Del Sordo and Thomas Mormann

1. *Introduction*

In January 1923 Carnap completed a manuscript with the lengthy title *Die Quasizerlegung. Ein Verfahren zur Ordnung nichthomogener Mengen mit den Mitteln der Beziehungslehre* (henceforth *Quasizerlegung*).[1] *Quasizerlegung* can be considered, together with a few other manuscripts of the early 1920s, as important groundwork of Carnap's first opus magnum *Der logische Aufbau der Welt* (1928; Eng tr. 1967, henceforth *Aufbau*) (see Proust 1986, Eng.tr. 1989: section 4; Mormann 2009). The content and text of *Quasizerlegung* can be easily reckoned as a theoretical base for many passages of *Aufbau*. The treatment of quasi-analysis in the *Aufbau* and the distinction between property and relational descriptions (§§10, 71-74, *passim*) were, indeed, already examined in the previous typescript (1-2). Some examples of property and relational description (see for instance *Aufbau*: 20, 114-116) were also developed in detail there (1, 5). Similarly, the formalism that the author introduces in *Aufbau* (§97) and explains in the *Abriss der Logistik* (Carnap 1929) was already applied in the former work.

The history of the reception of *Quasizerlegung* is complicated. Although it has never been published "officially", it has attracted the attention of quite a few readers through the decades. Without claiming to be exhaustive, let us mention the following philosophically relevant episodes.

According to Carnap (1957), the contents of the manuscript were debated for the first time at the "Erlangen Conference" in March 1923 (cf. Del Sordo 2016: 205-6). As Damböck (2021: 23-7) emphasizes, the content of *Quasizerlegung*, together with those of other early works, conveys the core idea of what can be characterized as Carnap's *Herzensprojekt*, in accordance with a letter from Carnap to Franz Roh in 1925 (23). By this "*Herzensprojekt*",

[1] The unpublished manuscript is preserved at the Archives of Scientific Philosophy (ASP), Hillman Library, Carnap papers, University of Pittsburgh (RC-081-04-01).

philinq X, 1-2022, pp. 231-254
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.401

he meant the proficient application of contemporary theory of mathematical structures to epistemological and ontological problems arising from German philosophy across the nineteenth and twentieth centuries (see Mormann 2016; Schnädelbach 1983, Eng.tr. 1984; cf. Damböck 2021: 21, 37-39). As Carnap notated in his diary (see Carus 2007: 157-158), his discussion partners in Erlangen, among them Hans Reichenbach, Kurt Lewin and Paul Hertz, did not understood very much the intention of his project. This course of misunderstanding has probably driven the author to abandon his *Herzensprojekt* at least in its original shape (cf. the conjecture of Damböck 2021: 27).

More than a decade later, *Quasizerlegung* became a topic in the correspondence between Carnap and Goodman dated January 1938 (see Proust 1986, Eng.tr.: 191-193). After Goodman's famous criticism of quasi-analysis (see Goodman 1951: ch. 5), Joelle Proust unearthed *Quasizerlegung* in 1986 in her book *Questions of Form.* She reconsiders the piece in its relevance to the *Aufbau* project and argues that Goodman's 1951 criticism rests on an underlying misinterpretation of the young Carnap's philosophical perspective (Proust 1986: Eng.tr.: 191-193). In this manoeuvre, Proust casts a renewed productive sidelight on the *Aufbau's* formal method by revealing that the axiomatic apparatus of quasi-analysis is much richer in *Quasizerlegung* than in *Aufbau*. While introducing his formal method in *Aufbau* (§80) Carnap mentions indeed only two of the four axioms applied in the manuscript (3). Elaborating Proust's philosophical investigations, Mormann in 1994 showed that axioms of *Quasizerlegung* can be interpreted as axioms for a structural representation. Elaborating this result, quasi-analysis can be aligned therefore with some of the main algebraic results of the 20th century mathematics (see Mormann 2009: 277 *passim*; Davey *et al.* 2002: chs. 5, 11). By means of further historico-philosophical investigations Mormann (2016: especially 118-129) also traces the origin of Carnap's quasi-analysis back to the German cultural *milieu* of *Lebensphilosophie* and in particular to the philosophy of neutral monism that authors like Mach, Ziehen and Avenarius among others developed around 1900.

A minor obstacle for the contemporary reader's understanding of *Quasizerlegung* (and more generally of the more formal passages of the *Aufbau* and its significance) resides in the fact that Carnap used in these texts an outdated formalism of logic and the theory of relations that is essentially that of Russell/Whitehead's *Principia Mathematica.* Actually, Carnap used not more than the most elementary terminology of propositional logic and theory of relations that can easily be translated into the nowadays more familiar set-theoretical terminology. Moreover, the "theorems" of *Quasizerlegung* are almost always logically and mathematically rather trivial reformulations of the definitions and need no

more than few lines to be proved.[2]

A short presentation of the symbolism employed in *Quasizerlegung* can be found in his booklet *Abriss der Logistik. Mit besonderer Berücksichtigung der Relationstheorie und ihrer Anwendungen* (Carnap 1929, henceforth *Abriss*).[3] In the *Aufbau* (§97) one can find a short list of the terminology used in this work, *Quasizerlegung* and other early writings of the author. In sum, the contemporary reader should have no unsurmountable difficulties to translate all formulas that Carnap used in these works in the more familiar terminology of informal set-theory. Be this as it may, in order to render *Quasizerlegung* more easily accessible for the contemporary reader we have added an appendix containing some (hopefully) useful hints and explanations that should make reading the manuscript more easily.

After these preliminary remarks let us now come to a crucial point, namely, a compelling argument why – after all – *Quasizerlegung* deserves to be carefully studied. In order to make plausible the claim that the manuscript has more to offer than an ingenuous formalism without philosophical significance, it is necessary to show that its perspective can help for a better understanding of significant concepts of contemporary philosophy and science. This is exactly what we want to sketch in the following. More precisely, we claim that the method of Carnapian quasi-analysis (as presented in its most elaborated form in *Quasizerlegung*) may be understood as a prototype of a promising mathematical philosophy in the sense that recently was explicated by Leitgeb (2012). Mathematical philosophy in this sense can be traced back to Russell's trail-blazing *Our Knowledge of the External World as a Field for Scientific Method in Philosophy* (Russel1 1914). We'd like to put forward the thesis that mathematical philosophy constitutes a current research field that closely inherits the original spirit of *Quasizerlegung* and Carnap's *Herzensprojekt* program in general. Indeed, mathematical philosophy promotes a view according to which philosophy is neither an ancillary discipline accompanying science, as it is in the analytic approach of logical empiricism, nor a part of science itself, as in the naturalized epistemology program of Quine and others (Leitgeb 2012: 267-268). Quite the contrary: mathematical philosophy ostensibly exhibits a close affinity with the idea of *Quasizerlegung* and Carnap's *Herzensprojekt* in general, by pursuing a philosophical research through mathematical, logical, and scientific methods and maintaining at the same

[2] In modern theory of relations there are non-trivial contentful theorems (see Maddux 1991; 2006; Givant 2017). Some interesting theorems concerning quasi-analysis and the complexity of similarity structures have been proved by Brockhaus 1963. They are discussed in Mormann 2009.

[3] Since some years a free electronic copy of *Abriss* in pdf is available in the internet.

time philosophy as a discipline in its own right, possessing its own problem, concepts and history (*ibid*.: 268-269).

The fundamental significance of *Quasizerlegung* for mathematical philosophy has been largely ignored so far. Either the early Carnap is conceived as a (proto-)analytical philosopher who achieved his philosophical maturity only in his later works, such as *Logical Syntax of Language* or *Empiricism, Semantics, and Ontology*. Or, this point of view is preferred by more historically inclined philosophical spirits, the early Carnap is conceived as a somewhat peculiar neo-Kantian philosopher. As we want to show both accounts fail to meet the full significance of *Quasizerlegung* for Carnap's philosophy in particular, and for modern scientifically-minded philosophy in general. Our proposal is, instead, to conceive the manuscript as the prototype of a scientifically-minded mathematical philosophy where two strands of thought come together, namely, the theory of representation, from working mathematics, and neutral monism, from philosophico-scientific research. The general target of our investigation is the theoretical significance of the recent historico-philosophical discoveries according to which quasi-analysis originates as both representation theorem (Mormann 2008) and means to reconcile the dichotomy between *Leben* and *Geist* sprouted from the soil of the early 20th century German thought in *Lebensphilosophie* and of neutral monism in particular (Mormann 2016). Indeed, given that quasi-analysis turns out to be a representation theorem, what does such a theorem serve for in neutral monism? How can it emerge over the course of a philosophy that, like neutral monism, elaborates ideas and tendencies that traditionally drift far apart from the analytic leanings of Carnap's later thought (cf. Schnädelbach 1983, Eng.tr.: ch. 5)? Now, formulating a definite and comprehensive answer to these philosophical questions is a profound task of inquiry and necessarily exceeds the scope of a single paper. Our particular aim is more modest consisting in paving the way for this kind of answers. In order to contribute to the reception of *Quasizerlegung* in this sense, the structure of this paper is as follows. In section 2, we specify some reasons for considering the philosophical relationship between neutral monism and representation theorems as a difficult, but highly rewarding issue of research. In section 3, we pretend to open the way for such an account by uncovering relevant points of convergence between the philosophical and mathematical enterprises of neutral monism and representations. In section 4 we conclude with some general remarks on the significance of *Quasizerlegung* for Carnap's philosophy and scientifically minded contemporary philosophy in general. In the appendix, we briefly explain the formalism that Carnap employed in the manuscript, in particular the formal concept of relation and the basic notions of the calculus of relations.

2. *Early Carnap's project: bringing together representation and neutral monism*

In order to argue for the relevance and possible fruitfulness of early Carnap's project of a mathematical philosophy (realized only in a preliminary and incomplete form in *Quasizerlegung* and other early manuscripts) let us comment upon some pieces of Carnap's earliest philosophical production of the 1920s. Admittedly, a lot of guesswork and speculation is involved in this endeavour. An essential ingredient of Carnap's project was geometry in a general sense. More precisely, geometry understood as synthetic geometry as a general theory of "*Ordnungsgefüge"*(cf. Mormann 2003: 47-50). Indeed, in *Quasizerlegung*, in *Der Raum*[4] (Carnap 1922b: ch. I, The Formal Space), and in the *Aufbau* (§70) Carnap used the very same example of color stripes and their similarity relations as an argument to argue for a geometry as a general (even universal) representational theory of order.

The concept of representation has, persistently, maintained a central position in the history of philosophy. Consequently, it has become a highly ambiguous concept with many different, even inconsistent meanings. In order to forestall any unnecessary misunderstandings, we want to point out from the outset that we subscribe to a monistic concept of representation that emphasizes the unity of the representational realm.

As the Neo-Kantian Ernst Cassirer (a frequent critical reference point of many logical empiricist such as Schlick, Frank, and Carnap in the first third of 20th century) has pointed out, many metaphysical doctrines tend to separate the domains of the representing and the represented, often conceived as the domain of "thought" ("*Denken*") and the domain of "things" ("*Dinge*") (cf. Cassirer 1910: 359). Thereby different "natures" are ascribed to both domains leading to the well-known riddle of how human knowledge is able to bridge the abyss between the allegedly totally separated two domains. In this sense, also in Cassirer's Neo-Kantianism epistemology an element of neutral monism can be identified.[5]

[4] In *Der Raum* (Carnap 1922b) Carnap explicitly mentioned the conceptual affinity of quasi-analysis with synthetic geometry: He pointed out that a classical theorem of Desargues may be understood as a quasi-analytical representation. A direct reference to geometrical representation theory is, however, neither mentioned in *Quasizerlegung* nor in the *Aufbau*. Thus, up to now, the philosophically crucial connection between geometry, order theory, quasi-analysis, and constitutional theory of the *Aufbau* has been rather ignored by Carnap scholars.

[5] The monist thesis of "representation first" may be backed also anthropologically. According to Ian Hacking: "The first peculiarly human invention is representation. Once there is a practice of representing, a second-order concept follows in train. This is the concept of reality, a concept which has content only when there are first-order representations. It will be protested that reality, or the

In order to better understand to what the primacy of representation amounts it is expedient to have a more detailed look on what a successful representational practice is expected to offer. We'd like to put forward the thesis that representations in mathematics and other cognitive enterprises aim at representation theorems. This is of critical importance since representation theorems can be characterized as monistic representations that assume the represented objects as primitive entities. The kind of representations that feature in mathematical representation theorems is, however, somehow unusual from both the analytical and ordinary points of view. Indeed, ordinary and analytical representations – one may think of representation as it arises, for example, in the philosophy of the first Wittgenstein – assume the represented object not to be primitive, but rather embedded in a domain of other objects that may serve in practice as representing props. As we shall see in a moment, representation theorems behave quite differently. They ferret out, in fact, representing props only via exploration of the represented object, since other domains of objects fall out of reach, by assuming primitivity.

What are representation theorems and what are they good for? A valuable attempt to answer these questions in an accessible, but informative and substantial way has been made by Davey and Priestley in their textbook on lattices and order (Davey *et al.* 2002). Even if they only deal explicitly with a special class of ordered structures, namely, lattices. In fact, their arguments apply to a much wider class of representations and representation theorems.

To be specific, a representation theorem for a class of lattices aims at a better theoretical and practical understanding of a class L of lattices. This is to be achieved by finding for the members of L sets P of basic building blocks ("atoms", "prime elements", "irreducible elements" etc.). These building blocks are either elements of the lattices or generated by them.

With respect to a lattice L, its "generating" set P has to fulfil (in some sense to be specified) the following requirements (*ibid.*: 112):

(A) The elements of P are readily identifiable. The cardinality of P should be as small as possible;

(B) The ordered set P should determine L in a unique way;

(C) The construction of L starting from P should be executable in a simple way.

The conditions (A) – (C) are to be interpreted as general guidelines or blueprints for conceptual constructions. They can be carried out and evaluated in various ways. For instance, with respect to (A) in many cases it is not uniquely

world, was there before any representation …. Of course. But conceptualizing it as reality is secondary" (Hacking 1983: 136; cf. also Rheinberger 2010: ch. 6).

determined what entities have to be chosen as appropriate "building blocks". Usually, it is not at all obvious what objects have to be chosen as elements of P. Only in rare cases it is obvious what the "atoms" are to be.

To be specific, let us consider some examples. For instance, in the case of an "atomic Boolean algebra" B it is rather clear that the elements of its generating set are to be taken as its atoms (smallest non-zero elements). However, already for non-finite Boolean algebras it is no longer possible to assume the existence of atoms in the usual sense. In general, for non-Boolean lattices atoms may not be available. Instead, appropriately chosen structures such as prime ideals have to be found that can play the role of building blocks. This task may require a high degree of ingenuity and technical skill. This is shown, in a particularly impressive manner, by Stone's famous Representation theorem that is to be considered as a paradigmatic case of a representation theorem *überhaupt*.

As Davey *et al*. also point out (2002), the requirements (A)-(C) have to be evaluated in a flexible way. For instance, in the case of Birkhoff's representation theorem, the requirement (A) of "smallness" is clearly satisfied (*ibid*.: 121). On the other hand, by taking Stone's representation theorem, "smallness" has to be evaluated with a grain of salt. In particular, when L is an infinite complete Boolean algebra, P is anything but small. Indeed, by the theorem of Balkar-Franěk (see Koppelberg 1989: 196), the cardinality of P is equal to the cardinality of L. Nevertheless, also in this case, the constitution of L from P is to be considered as an important conceptual achievement for other reasons.

Further problems arise in interpreting (A) as a guiding principle from the "easy identifiability" condition on P. Surely, one can think of this condition as satisfied by building block structures in Birkhoff theorem and finite Boolean algebra representation (Davey *et al*. 2002: 116-121). Its fulfilment becomes seriously debatable, however, whenever one has to appeal to the axiom of choice, or other maximality principles, to prove the existence of prime ideals or prime filters as is the case for Stone's representation theorem and many other modern theorems of this kind. This situation occurs in lattice theory, in the case of Stone and Priestley representation theorems (*ibid*.: chs. 10-11), as well as quasi-analysis (Mormann 2009: 259).

Let us now consider requirement (B). The theorems of Birkhoff and of finite Boolean algebras undoubtedly meet it (Davey *et al*. 2002: 114-116). To render (B) more precise, one should probably strengthen its criterion by also adding fundamental relations other than the order one. Let us finally consider (C). As Davey *et al*. (2002) emphasises, one cannot consider it as overall satisfied even in the realm of lattice theory. A general representation theorem of finite complete lattices would hardly meet it, indeed (*ibid*.: 168).

If one considers representation theorems from a wider philosophical point of view, then it might be convenient to interpret (C) in terms of an epistemic economy principle of some kind. In this case, since the fourth axiom of quasi-analysis too (3) is interpreted as an economy requirement (Mormann 2009: 253), the same axiom can be considered as an instance of (C). Accordingly, just as in lattice theory, (C) is not always satisfied by quasi-analysis either, as some additional conditions might conflict with it (*ibid*.: 262, *passim*).

Taking into account the difficulty of finding appropriate building blocks for many apparently simple mathematical representations, it might not be surprising that the analogous task of finding appropriate "neutral elements" in the realm of philosophy in general features a similar difficulty. Neutral monism is a comprehensive philosophy which includes both epistemological and ontological theses. According to it, the world we live in is entirely constituted by systems of "neutral" elements. To borrow a famous (or notorious) Neo-Kantian pun the neutral elements are not "*gegeben*" to us, rather, the task of finding them is "*aufgegeben*".

Neutral elements are identified by being structural, qualitative and pre-cognitive entities (see Del Sordo 2021: ch. 2). Among these features, that of "being pre-cognitive" is allegedly the most puzzling. Pre-cognitive nature renders, indeed, neutral elements elusive to any form of cognitive attitude purportedly focused on them. Because of this, the matching of our common believes to the theses of neutral monism is not an easy task to carry out, as Textor (2021: 33-37) recently showed. Such an epistemic difficulty has often led philosophers to weaken the strength of its theory. By virtue of the elusiveness of neutral elements, Tully (2003: 337-338), for instance, ends up reducing neutral monism to a metaphor or, at most, to a very abstract and formal hypothesis. Such interpretation does not provide, however, a coherent view of the movement. Indeed, it basically neglects the fact that its exponents were deeply engaged in finding appropriate theoretical strategy to overcome the epistemic elusiveness of their fundamental elements. Mach held, for example, that one day a future physiology would have empirically grasped what neutral elements essentially are (see Mach 1896, Eng. tr. 1898: 212; Banks 2003: 134). In this respect, Russell too seemed to share, at least in some passages, a line of thought allied with Mach's (see for example Russell 1927: 281-282).[6]

If we confine our discussion to the standard authors of neutral monism, i.e., Ernst Mach, William James and Bertrand Russell, then the syntheses of

[6] Insofar as its content is currently under debate, we must be cautious in making this statement. Actually, in the scholarship of neutral monism (see for example Wishon 2021: 139-141), Russell's neutral entities afford also a reading in terms of inscrutables, which renders them ungraspable by any form of knowledge.

the movement proposed by Banks 2014 and Stubenberg 2016 can be arguably considered as a unitary meta-theoretical account of the movement. Both have the drawback, however, not to put neutral monism in a wide enough historico-philosophical context. Fundamental elements with the same characteristics of the neutral ones, i.e., being qualitative, structural and subjectless, can be encountered indeed in the philosophical perspectives of Husserl's phenomenology and Bergson's metaphysics, to mention just two (see Schnädelbach 1983: Eng.tr., ch.5, also 148, *passim*). Husserl and Bergson, which one may eventually qualify as non-standard authors of neutral monism, were not content, just as Mach and probably Russell, with merely abstract or metaphorical proposals. They rather developed sophisticated methodologies, respectively based on *epoché* and intuition, to cognitively grasp the essence of their neutral fundamental entities (see for example Husserl 1913: § 63; Bergson 1911: ch. 3).[7] After all, an early version of quasi-analysis may be found in the work of Ziehen (1913: 3; cf. also Mormann 2016: 116) as a suitable philosophical method to address an epistemic challenge that was analogous to the one addressed by Mach, Husserl or Bergson. Quasi-analysis was meant, indeed, to free the constitution of reality in neutral monism from undue cognitive or epistemic assumptions (Ziehen 1913: 1-2, 177-178), a target that, according to Ziehen (1920: 217), phenomenological and scientific methods hinging upon extraordinary intuition or future physiology were not able to perform.

The considerations of this section entail that the theoretical meaning of the historico-philosophical origins of Carnap's quasi-analysis has to be explored by answering the questions: in what sense does quasi-analysis arise as an alternative philosophical method to those of other authors, and how can it eliminate undue epistemic assumptions within the constitution of reality in neutral monism? To answer these questions, plausible solutions to the above-mentioned meta-theoretical difficulties must be worked out. To this end, additional mathematical and philosophical topics should be deeply unfolded. Indeed, it is a matter of following the scientific development of representation theorems, perhaps using the formalism of category theory (for references on this see Davey *et al.* 2002), and the metaphysics and epistemology of order as it was developed around 1900 (for references and insights on this issue see Ziche 2016). Delving into these questions lies beyond the scope of this paper. Even so, we hope this section has exposed how, in spite of its difficulties, a meta-theoretical examination of neutral monism, representation theorems and their partnership may be a highly rewarding research topic.

[7] For a detailed study of the general connection of Husserl and Bergson with neutral monism their most relevant works are *Ideas 2*, *Analysis Concerning Passive and Active Synthesis* and *Matter and Memory*.

3. *Neutral monism and representationalism: towards a common program*

Some basic points of convergence between neutral monism and mathematical representation are undeniable. First of all, they both carry out a three-part theoretical program comprising *pars destruens* and *pars construens*. They start with complex relational entities, like the natural world or abstract algebras, whose constitution they want to clarify. To gain a better understanding of such complex entities, neutral monism and mathematical representation reduce them to building block structures (*pars destruens*). Forgetting[8] provisionally the information one has about the entities in their unreduced form, they finally elaborate perspicuous reconstructions of them by using only information provided by the building block structures (*pars construens*). This three-part program can be condensed to the following schema:

	Complex Structural Entity	Building Block Structures *(pars destruens)*	Perspicuous Reconstruction *(pars construens)*
NEUTRAL MONISM	Natural World	Neutral Elements	Neutral Elements (Perspicuous) Natural World
MATHEMATICAL REPRESENTATION	Abstract Structures	Prime Elements	Prime Elements (Perspicuous) Abstract Structures

Within this schema, one can track down three additional patterns of affinity between neutral monism and mathematical representation.

1. The items, e.g. everyday natural objects or algebraic elements, which the complex entity consists of, turn often out to be systems of building block elements (whether neutral or prime) (cf. Banks 2014: ch. 1; Davey *et al.* 2002: chs. 5, 11).
2. In mathematical representation building blocks tends to be relation-

[8] The term "forgetting" has been chosen on purpose here. "Forgetful functors" are a basic concept of category theory that may be considered as a generalization of lattice theory playing a prominent role in the foundations of mathematics, informatics, and theoretical computer science (see for instance Simmons 2011: 76). "Forgetful functors" apply to those structures whose relations or operations must be set aside. Even for philosophers with only a rudimentary education in history of philosophy it is impossible not to detect an epistemological affinity between the operations of 'forgetful functors' and '*epoché*'. They both act, indeed, in such a way as to put some previously acquired knowledge into brackets.

ally, or structurally, poorer than the initial complex entities (see for example *ibid*.: 121, 262). Hence, we obtain that: if one epistemologically assumes that knowing is a matter of connecting, linking or ordering entities (see for example Ziche 2016: 91, *passim*), then the relational poorness of prime elements might simulate, or approximate, a pre-cognitive condition of sorts. If this is right, then one may also argue that the building block structure of quasi-analysis originally serves the purpose in neutral monism of emulating the elusive pre-cognitive condition of neutral elements. This hypothesis is, however, momentarily difficult to ascertain or generalize. From the historico-philosophical point of view, its ascertainment requires further examinations that can be unfolded better in papers *ad hoc*. Also, from the mathematical point of view, it cannot be generalized either, for the problem of determining the building block structures, i.e., problem of representability (see Davey *et al*. 2002: 261), still remains unsolved in many cases.

3. In the *pars construens*, both mathematical representation and neutral monism must comply with an economy requirement of some kind. Concerning mathematical representation, we have already seen in section 2 an economy requirement showing up both in (C) and in the fourth axiom of *Quasizerlegung* (3). Concerning neutral monism, such requirements have been applied at least in the form of Mach's normative economy of thought (Banks 2004: 24-5), where natural world is taken to be a parsimonious epistemic construct underpinned by an ontological array of pre-cognitive and chaotic neutral elements.

The above-considered three-column schema and points 1-3 do not contain a full-fledged exploration of the relationship between neutral monism and representation theorems. Nevertheless, they can be used as an entering wedge to further understand the philosophical origin of Carnap's quasi-analysis and envisage applications of mathematical representation in philosophical projects encompassing, like neutral monism, both ontology and epistemology.

4. *Concluding remarks*

Quasizerlegung is a piece of philosophy that defies straightforward classification. On the one hand, its formal aspects led some commentators to classify it as a sample of (early) analytic philosophy. If analytic philosophy is characterized, however, as the philosophical current according to which a philosophical account of thought can be attained only through language (see Dummett 1994: 4), then *Quasizerlegung ipso facto* does not appear to be very analytic. Indeed, the text features an amalgam of heavily loaden philosophical concepts, connected with metaphysical irrationality and ontological neutrality. On the

other hand, it would be too simple to interpret it as a work of neo-Kantianism or Husserlian phenomenology. Mormann (2006: 27-33 ff.) and Damböck (2021: 39-41) have shown that interpretations of this kind, for instance Friedman (2000) and Mayer (2016), overlook many essential features of early Carnap's work.

In sum, interpretations that force young Carnap's work into ready-made categories either of analytic philosophy, neo-Kantianism, or phenomenology turn out to be Procrustean beds for this text. This nourishes the suspicion that classical philosophical categories are too rigid to capture *Quasizerlegung*'s true meaning. One may conjecture that other less known and subterranean traditions are at stake here – one may think for example of the so-called "lost" neo-Kantian tradition (Beiser 2014) of Herbart, whose influence on neutral monism only recently has been re-appreciated (see Banks 2003: ch. 3)

Until today, some authors, although engaged in Carnapian scholarship, simply ignore the concept of quasi-analysis. For instance, in Carus influential book, *Carnap and Twentieth-Century Thought. Explication as Enlightenment* (Carus 2007) the concept of quasi-analysis does not even appear once, although this book claims to deal with the significance of Carnap's philosophy in general. Chalmers too, who, in his bulky *Constructing the World* (2012), explicitly pretends to resuscitate Carnap's "Aufbau-program", but does not mention the method of quasi-analysis at all.

In other publications dealing with Carnap and Carnapian philosophy quasi-analysis scores better and pops up quite often. Nevertheless, its role is usually restricted to a sort of philosophical Cinderella. The concept is briefly mentioned, but almost never treated in detail. In most chapters of Damböck *et al.* (2021) anthology, *Der junge Carnap in historischem Kontext 1918 - 1935*, quasi-analysis either is not mentioned or, when it is mentioned, is introduced *ex abrupto* without any formal or informal explanation of what it is about or is not. This treatment of a philosophical Cinderella was, so to speak, familiar to *Quasizerlegung* from the very beginning. In fact, as we have already seen, Carnap himself abandons the content of the manuscript together with his *Herzensprojekt* after being misunderstood by colleagues in Erlangen and advised by Reichenbach to bring the philosophical focus away from the overly general attitude (according to Reichenbach) of his early *Konstitutionstheorie* (see Damböck 2021: 25-26).

In sum, the history of *Quasizerlegung* and, more generally, of quasi-analysis as subjects of scientific and philosophical research has not been a lucky one. Thus, *Quasizerlegung* may be considered as a kind of Kuhnian anomaly in the history of epistemology and philosophy of science that has defied philosophical paradigm, be it analytical philosophy proper, logical empiricism, or main-

stream philosophy of science. In our opinion, the idea of quasi-analysis has to be seen as a philosophical challenge which is able to blow up the traditional borders of philosophical research. This is all the more true as quasi-analysis, in virtue of its application of mathematical methods and leanings towards problems of *Lebensphilophie* and neutral monism in particular, is a probably unstable mixture of scientifically minded philosophy and irrationalist metaphysical tendencies. This holds, in particular, since mathematics and representation are "protean" concepts that are realized in many different and varying ways (see Mac Lane 1986). The same mathematical structure may have, in fact, many empirical realizations, and representations, in turn, combine technical complexity and overall applicability in a host of different and allegedly divergent scientific and informal contexts. Neutral monism, on its side, runs afoul of philosophies that neatly separate mind and matter as many philosophers from Descartes onwards have done. It puts forward the ingenious hypothesis of bridging the gulf between mind and matter under the aegis of elusive, but still empirically effective, "neutral" entities.

The individual fruitfulness of mathematical representationalism and neutral monism is difficult to be overlooked and underestimated. But what about the Carnapian project of combining them in quasi-analytical framework? As already explained, this is a huge and complex question to answer. We can reasonably claim, however, that the multifaceted character of representation plays a pivotal role here. The accuracy of this submission depends, however, on the idea of representation one subscribes to. Indeed, if one assumes representation as "kopeyliche Betrachtung" that was criticized already by Kant (see Mormann 2018: 3), the marriage of mathematical representation and neutral monism is bound to end in an unhappy and fruitless relation, since the array of entities that the latter needs to represent are epistemically fleeing, or elusive, and thus cognitively unavailable to be copied. On the other hand, things may look brighter, when one shifts to a wider and more flexible account of representation where representing is not a matter of copying but, in a modernized Kantian-style, one of historically intervening and constituting both scientific and ordinary objects (*ibid.*: 5). In this sense, a constitutive and monistic view of mathematical representation may be a good candidate for overcoming restrictive epistemic dichotomies, like the ones already encountered of mind vs. matter and *Leben* vs. *Geist*.

Appendix

The formalism of *Quasizerlegung*: some hints and explanations for the contemporary reader

The formalism of *Quasizerlegung* harks back to the calculus of relation that Augustus De Morgan, Charles Sanders Peirce and Ernst Schröder developed in the second half of the 19th Century (see Givant 2017: 27-8[9]). Aside from theorems (1)-(7) that are formulated in natural language (3-4), it is evidently applied in the discussion of theorems (*Lehrsätze*) (8)-(47) and also invoked in the first section of the manuscript (2), where the author claims to justify an overall applicability of his method, regardless of the basic relations that one assumes at the outset. Let us refer, for simplicity, to binary relations and define them as follows (cf. *Abriss*: 25):

> (D1) A binary relation R on a set X is defined to be a subset of the set $X \times X$ of all ordered pairs $\langle x_1, x_2 \rangle$ of elements x_1, x_2 in X.

There are different ways of visualizing relations. Indeed, one may visualize them by using graphs, matrixes or lists of ordered n-tuples (see *Abriss*: 26-8; Givant 2017: ch. 1, *passim*). As one can clearly see in the text (5 ff.), in *Quasizerlegung* Carnap chose the third way and introduced as an example a similarity structure of the 12 sounds h, l, k, … by listing its positive pairs.

A relation R is said to be *included* in a relation V if and only if every pair of R belongs to V. Following *Abriss* (28) and *Quasizerlegung* (2), inclusion may be symbolically expressed by writing,

$$V \subset R\,.$$

Relational inclusion corresponds to set-theoretical inclusion, $\subseteq$, (cf. Givant 2017: 2) and satisfies the laws of reflexivity, anti-symmetry and transitivity (see *Abriss*: 29). Also, two relations R and S are defined to be *equal* $V = R$ if and only if both $V \subset R$ and $R \subset V$ hold, i.e., if and only if they contain the same ordered pairs. Two special relations are introduced, namely the *identity*

[9] For further historical details see Maddux 2006: 1, *passim*; Maddux 1991.

relation (2), I, and the *diversity relation* (*Abriss*: 26), ≠. The former is a reflexive, symmetric and transitive relation and consists of pairs of equal elements. The latter is a symmetric relation and consists of pairs of unequal elements.

There are several constructs for building new relations from already given ones. Suppose that R and S are again relations. The *union* $R \dot{\cup} V$ of *R* and *V* (2; *Abriss*: 28-9) is the relation consisting of the pairs that are either in R or in S,

$$R \dot{\cup} V = \{\langle x, y\rangle | \langle x, y\rangle \in R \text{ or } \langle x, y\rangle \in V\}$$

The *intersection* $R \dot{\cap} V$ of *R* and *V* (2 ff.) is the relation consisting of the pairs that are in both *R* and *V*,

$$R \dot{\cap} V = \{\langle x, y\rangle | \langle x, y\rangle \in R \text{ and } \langle x, y\rangle \in V\}$$

If *R* is a relation on a set *X*, then the *complement* of R, $\dot{-} R$ (*Abriss*: 28), is the relation consisting of the pairs that are in $X \times X$, but not in *R*. Also, the *difference* of *V* and *R* (*ibid.*) is the relation consisting of the pairs that are in *R*, but not in *V*,

$$R \dot{-} V = \{\langle x, y\rangle | \langle x, y\rangle \in R \text{ and } \langle x, y\rangle \notin V\}$$

Union, intersection, difference and complement of relations basically correspond to set-theoretical union, ∪, intersection, ∩, and difference, −, and satisfy the related laws of the set-theoretical operations, e.g., associativity, commutativity, distributivity, De Morgan etc. (*ibid.*).

Finally, the *converse*, or *inverse*, of *R*, $\breve{R}$, (5,6) consists of the pairs in *R*, but with the reversed order, in symbols,

$$\breve{R} = \{\langle x, y\rangle | \langle y, x\rangle \in R\}$$

It satisfies the law of involution $\breve{\breve{R}} = R$, i.e., the converse of the converse of *R* is equal to *R* itself (*Abriss*: 36; Cf. also Maddux 2006: 21).

While the constructs above build new relations from already given ones, the following ones serve, instead, to decompose relations. The *domain* D'*R* of the relation *R* consists of all the left-hand members of the pairs in *R* (7). In symbols,

$$\dot{\supset}\,'R = \{x | \text{ for some } y \langle x, y\rangle \in R\}$$

Conversely, the *range*, $\dot{\subset}\,'R$, consists of all the right-hand members of the pairs in R,

$$\dot{\subset}\,'R = \{y | \text{ for some } x \langle x, y\rangle \in R\}$$

The *field* C'*R* is the set-theoretical union of the two previous sets, C'*R*= ⊐'*R* ∪ ⊏ '*R*. For further information about the formal properties of these operators and alternative symbolic expressions, consult *Abriss* (35-8) and Givant 2017 (14, 148). Among the methods of decomposing constructions, one may also characterize the *projection* $\vec{R}'y$ which maps each element of the field C'*R* to the set of its left-side companions in *R*. In symbols,

$$\vec{R}'y = \{x|\langle x, y\rangle \in R\}$$

Mutatis mutandis, one can also characterize the projection $\overleftarrow{R}'x$ (see *Abriss*: 35). Projections appear very often throughout the manuscript (5-6 ff.). More specifically, Carnap uses their respective specular operators to define the concept of similarity neighbourhood (5). Obviously, whenever the relation *R* is, such as similarity, a symmetric relation (*ibid.*), $\vec{R}'x$ and $\overleftarrow{R}'$x result in the same outputs. As a consequence of this, while choosing to work with similarity as basic relation, the author does not tell the difference between the two operators in the text.

Projections can be also defined as functions that map C'*R* to its powerset. Following the calculus of classes, both in *Abriss* (25) and *Quasizerlegung* (10) Carnap indicates the powerset of a set *X* by Cl'*X*. Abstraction operators and union/intersection of families of sets are correspondingly expressed by capped variables without brackets followed by abstraction conditions, $\hat{x}$ (... x ...) (cf. for instance 5, 7, definitions (8) and (16)) and apostrophized letters " s' ", " p' " followed by the family of sets they are wanted to unify or intersect (cf. for instance 9, 10, theorems (25) and (34)). These notations are quite different from the more recent one, {x| ... x ...}, ⋃, ⋂, and go back to *Principia Mathematica* (see also Marciszewski 1981).

Two further methods of constructions can now be introduced, namely *relative product* and *restriction*.

The *relative product*, or *(relational) composition*, *R*|*V* (2; *Abriss*: 38-9) is the relation consisting of the pairs $\langle x, z\rangle$ such that for some y if $\langle x, y\rangle \in R$ and $\langle y, z\rangle \in V$, then $\langle x, z\rangle \in R|V$ (see also Maddux 2006: 7 and Givant 2017: 6). It satisfies the law of associativity, i.e., (*R*|*V*)|*T* = *R*|(*V*|*T*) (see *Abriss*: 39 for instance). The relational product of *R* with itself, *R*|*R*, is denoted by R^2. As is easily seen the relation *R* is a transitive relation if and only if $R^2 \subseteq R$ (see *ibid.*: 40).

In *Abriss* (37-8), there are several types of relation *restrictions*, like *domain-restriction* or *range-restriction*, expressed by operators ↾, 1. For simplicity, we will define only range-restriction, which is also the only one that *Quasizerlegung* applies. In particular, if $\beta \subseteq \sqsubset' R$, then $R \upharpoonright \beta$ is the set of the *R*-pairs whose right-hand members are contained in β. In symbols,

$$R \upharpoonright \beta = \{\langle x, y\rangle | \langle x, y\rangle \in R \text{ and } y \in \beta\}$$

Through restrictions, another special relation can be defined, which consists of the pairs of equal elements on the field of R. It is introduced in the first section of *Quasizerlegung*, where Carnap indicates it by $I \upharpoonright \mathrm{C}'R$ (2). The same relation can also be expressed by R^0 (see *Abriss*: 37-8).

Finally let us mention the *iterated relative product* of a relation with itself (2). It is indicated by R_* and consists of the following union of relations:

$$R_* = R^0 \dot{\cup} R \dot{\cup} R^2 \dot{\cup} R^3 \dot{\cup} \ldots$$

R_* is defined to be a *R-chain* if, whenever $< x, y >, < y, z > \in R_*$, we have $< x, z > \in R_*$ for all $\mathrm{x, y, z} \in \mathrm{C}'R$. In case *R* is infinite one seems to need the axiom of choice or some similar principle to ensure the existence of R_*. For further formal properties of *R*-chains, consults *Abriss* (56-7).

It might be useful to conclude this section by briefly clarifying the meanings of the logical and auxiliary symbols that are applied in the manuscript. Concerning the logical symbols, Carnap uses the symbols $\supset, ., \vee, -, \equiv$ to respectively indicate connectives $\rightarrow, \wedge, \vee, \neg, \leftrightarrow$. Moreover, the operators $\exists x, (x)$ and $\imath$ stand for existential, universal and definite description quantifiers (see *Abriss*: chs. 4-7). Concerning the auxiliary symbols, following PM, Carnap used a kind of "point calculus" for round, square, and curly brackets that use to be employed in modern treatises. One can find the rules of their use in *Abriss* (9-10). Depending on the nesting level, more dots, e.g. :, ::, or :., stand for square or curly brackets. Dots lie beside connectives $\supset$, $\vee$ and $\equiv$, and follow both the symbol of deduction "$\vdash$" and quantifiers. Every point brings together symbols either to the end or to the next points.

We conclude this appendix putting together several definitions of "similarity circles" (6 ff.), or *Ähnlichkeitskreise*. Similarity circles are paramount in quasi-analysis, for they are the building blocks, or prime elements, (cf. § 3 above) of similarity structures[10]. They are also the nodal point of Goodman 1951 harsh criticism of unfaithfulness and inaccuracy against quasi-analysis (see Goodman 1951: 157-161; cf. Leitgeb 2007: 193-200 ff.). Carnap introduces similarity circles in *Quasizerlegung*, *Abriss* and *Aufbau*. Basically, their introductions display formal or informal language and differ in applying the constructs of relation algebra or not. For a set to be a similarity circle two fundamental conditions must be met, 1) that of being a homogeneous similarity structure (i.e. every element is similar to every other) and 2) that of being

[10] Some interesting non-trivial results of the theory of similarity circles can be found in Brokhaus (1963). Discussed in Mormann (2007).

trivially included in other similarity circles (i.e. if A and B are similarity circles and $A \subseteq B$, then $A = B$). Before defining similarity circles, the author introduces similarity relations. To this end, he explicitly employs the above-mentioned constructs of relation algebra only in *Aufbau*, where similarity, S, is defined by applying the relational operators of union $\dot{\cup}$, conversion $\breve{}$ and identity ^{0} to a given asymmetric relation R: $S = R \dot{\cup} \breve{R} \dot{\cup} R^0$ (see *Aufbau*: 179). On the other hand, *Quasizerlegung* (2), as well as Mormann (2009: 255), introduce similarity invoking symmetry and reflexivity axiomatically. Following *Aufbau* (§80), similarity circles are informally defined in the following way:

> (D2): A similarity circle, *SC*, is a subset of elements of a similarity structure such that: (i) any two elements of *SC* are similar one to the other; (ii) if an element is similar to all elements of *SC*, then it belongs to it.

The conditions (i) and (ii) express, respectively the above-mentioned informal conditions of homogeneity 1) and maximality 2). Formal definitions of similarity circle present different nuances in *Quasizerlegung* and *Abriss* in accordance with the author's choice of using the constructs of relation algebra or not. Let us begin with the definition provided by Mormann (2009: 259), which is more familiar to the reader:

> (D3) Let $(S, \sim)$ be a similarity structure. A subset T of S is a similarity circle (*Ähnlichkeitskreis*) of $(S, \sim)$ iff it is a maximal set of similar elements, i.e., iff it satisfies the following two conditions: (I) For all x, y $(x, y \in T \Rightarrow x \sim y)$; (II) For all $z \in S$ $(z \sim x$ for all $x \in T \Rightarrow z \in T)$.

As can be easily seen, conditions (I) and (II) of (D3) are but the formal version of conditions (i) and (ii) of (D2). As we shall see in a moment, (D3) strictly reflects the introduction of similarity circles in *Quasizerlegung* (6, Definition (13)), see below (D4). Differences between (D3) and (D4) lie in logical notations, calculus of classes with capped second order variables and punctuation (., :, :: etc.).

> (D4) We define the class of the <u>similarity circles</u> as follows: $sim = \hat{\beta}\,(x,y \in \beta \supset xSy :.(z):(u).u \in \beta \supset uSz.\supset.z \in \beta)$ Def.

The first and second parts of (D4), namely $x,y \in \beta \supset xSy$ and $(z):(u).u \in \beta \supset uSz.\supset.z \in \beta)$ respectively correspond to (i) and (ii). As can be appreciated, both definitions (D3) and (D4) employ predicative logic without resorting to the constructs of relation algebra. On the other hand, they clearly figure in the definition, here called (D5), given in *Abriss* (49):

(D5) SC = Df $\hat{\alpha}$ $(\alpha \uparrow \alpha \subseteq R :. (x): \alpha \subset \vec{R}'x \supset x \in \alpha)$

The definition (D5) uses constructs of relational algebra and assumes that *R* is a similarity relation. The first part of the definition, i.e., $\alpha \uparrow \alpha \subseteq R$, ensure *SC* satisfies the condition 1) of homogeneous similarity structure, while the second part, i.e., $(x): \alpha \subset \vec{R}'x \supset x \in \alpha)$, ensures 2), that is, *SC* to be maximal. The constructs of relation algebra here employed are those of double restriction, $\alpha \uparrow \alpha$ and projection, $\vec{R}'x$. The double restriction, which is not numbered among in the operators listed above, is defined in *Abriss* (38) and amounts to the product of α and itself, namely the set $\alpha \times \alpha$ of all ordered pairs $\langle x_1, x_2 \rangle$ of elements x_1, x_2 in α. Projection is used, instead, to define similarity neighbourhood in *Quasizerlegung* (5) and is of some importance in contemporary debate. As noticed by Carnap (6), constructing a quasi-analysis on the base of similarity neighbourhoods is not possible. They allow, however, as Mormann 2009 (269 ff.) has shown, to find new order structures from given similarity ones and, as a result, to defeat criticisms of similarity structure for being too weak to be useful (cf. Goodman 1951; Quine 1969). By the way, the possibility of generating an asymmetric relation from similarity through similarity neighbourhoods was already detected in *Quasizerlegung*, when Carnap applied the projection $\vec{S}'$ to define the relation xE_1y, which holds whenever the similarity neighbourhood of *x* is a subset of the neighbourhood of *y* (5).

Caterina Del Sordo
University of the Basque Country UPV/EHU, Spain
caterina.delsordo@ehu.eus

Thomas Mormann
Tsukuba, Japan
thomasarnold.mormann@gmail.com

References

Banks, Erik, 2003, *Ernst Mach's World Elements. A Study in Natural Philosophy*, Springer, Dortrecht.

Banks, Erik, 2004, "The Philosophical Root of Ernst Mach's Economy of Thought", *Synthese* 138, 1: 23-53.

Banks, Erik, 2014, *The realistic Empiricism of Mach, James and Russell. Neutral Monism Reconceived*, Cambridge University Press, Cambridge.

Beaney, Michael, 2004, "Carnap's Conception of Explication: From Frege to Husserl?", in Awodey, Steve, and Klein, Carsten ed., *Carnap Brought Home: The View*

From Jena, Open Court, Chicago.

Beiser, Frederick, 2014, *The Genesis of Neo-Kantianism, 1796-1880*, Oxford University Press, Oxford.

Brockhaus, Klaus, 1963, *Untersuchungen zu Carnaps Logischem Aufbau der Welt,* Unpublished Dissertation, Universität Münster, Germany.

Bergson, Henry, 1907, *L'Évolution créatrice*, Félix Alcan, Paris; Eng. tr. by Arthur Mitchell 1998, *Creative Evolution,* Dover, New York.

Cassirer, Ernst, 1907, "Kant und die moderne Mathematik", *Kant-Studien* 12: 1-49.

Cassirer, Ernst, 1910, *Substanzbegriff und Funktionsbegriff*, Wissenschaftliche Buchgesellschaft, Darmstadt.

Carnap, Rudolf, 1922a, *Vom Chaos zur Wirklichkeit*, unpublished manuscript RC-081-05-01, University of Pittsburgh.

Carnap, Rudolf, 1922b, "Der Raum. Ein Beitrag zur Wissenschaftslehre", *Kant-Studien*, Ergänzungsheft 56, Reuther und Reichard, Berlin.

Carnap, Rudolf, 1923, *Die Quasizerlegung – Ein Verfahren zur Ordnung nichthomogener Mengen mit den Mitteln der Beziehungslehre*, unpublished manuscript RC-081-04-01, University of Pittsburgh.

Carnap, Rudolf, 1928, *Der logische Aufbau der Welt*, Weltkreis, Berlin; Eng. tr. Rolf George 1967, *The logical structure of the World*, University of California Press, Berkeley/Los Angeles.

Carnap, Rudolf, 1929, *Abriss der Logistik*. Springer, Wien <https://archive.org/details/RudolfCarnapAbrissDerLogistik/page/n3/mode/2up>.

Carnap, Rudolf, 1957, *Unpublished autobiography*. UCLA Carnap Papers, Box 2, CM3: B29-B36.

Carus, André, 2007, *Carnap and Twentieth-Century Thought. Explication as Enlightenment*, University Press Cambridge, Cambridge.

Chalmers, David, 2012, *Constructing the World,* Oxford University Press, Oxford.

Damböck, Christian, Gereon Wolters, 2021, eds. *Der junge Carnap in historischem Kontext: 1918 – 1935/Young Carnap in an Historical Context: 1918 – 1935*, Veröffentlichungen des Instituts Wiener Kreis, Springer, Wien.

Damböck, Christian, 2021, "Die Entwicklung von Carnaps Aufbau 1920–1928", in Damböck, Christian, and Gereon Wolters eds., *Der junge Carnap in historischem Kontext: 1918-1935*: 19-53

Davey, Brian, Hilary Priestley, 2002, *Introduction to Lattices and Order*, Cambridge University Press, Cambridge.

Del Sordo, Caterina, 2016, "Commentary. *Die Quasizerlegung. Ein Verfahren zur Ordnung nichthomogener Mengen mit den Mitteln der Beziehungslehre* by Rudolf Carnap", in Lanfredini, Roberta, Nicola Liberati, Andrea Pace Giannotta, Elena Pagni eds., *The Enactive Approach to Qualitative Ontology in Search of New Categories*, Humana.mente Journal of Philosophical Studies, 31,9: 205-210.

Del Sordo, Caterina, 2021, *Rappresentazione Strutturale delle Entità neutrali. Il Monismo Neutrale da un Punto di Vista Analitico e Storico-filosofico*, Ph.D. Dissertation, University of Urbino (Italy) / University of the Basque Country UPV/EHU (Spain).

Dummett, Micheal, 1993, *Origins of Analytical Philosophy,* Duckworth, London.

Givant, Steven, 2017, *Introduction to relation algebras*, Relation Algebra, 1, Springer.

Goodman, Nelson, 1951, *The structure of appearance*, Harvard University Press, Harvard.

Grosholz, Emily, 2007, *Representation and Productive Ambiguity in Mathematics and the Sciences*, Oxford University Press, Oxford.

Hacking, Ian, 1983, *Representing and Intervening, Introductory Topics in the Philosophy of Science*, Cambridge University Press, Cambridge,

Husserl, Edmund, 1913, *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie*, Max Niemeyer, Halle; Eng. tr. by Fred Kersten 1982, *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy—First Book: General Introduction to a Pure Phenomenology*, Nijhoff, The Hague

Koppelberg, Sabine, 1989, *Handbook of Boolean Algebras*, in Monk, Donald, Robert Bonnet eds. 1, North-Holland, Amsterdam and New York.

Lakoff, George, Rafael Núñez, 2000, *Where Mathematics comes from. How the Embodied Mind brings Mathematics into Being*, Basic Books, MIT press, New York.

Leitgeb, Hannes, 2007, "A New Analysis of Quasi-Analysis", *Journal of Philosophical Logic*: 181-226.

Leitgeb, Hannes, 2013, "Scientific Philosophy, Mathematical Philosophy and All That", *Metaphilosophy*, 44, 3: 267-275.

Mach, Ernst, 1896, *Populär-wissenschaftliche Vorlesungen,* Johann Ambrosius Barth, Leipzig; Eng. tr. by Thomas McCormack 1898, *Popular Scientific Lectures,* Open Court, Chicago

Mac Lane, Saunders, 1986, *Mathematics. Form and Function*, Springer, New York.

Mac Lane, Saunders, 1992, *The Protean Character of Mathematics*, in Echeverria, Javier, Andoni Ibarra, Thomas Mormann eds., *The Space of Mathematics. Philosophical, Epistemological, and Historical Explorations*, de Gruyter, Berlin: 3-13.

Maddux, Roger, 1991, "The origin of relation algebras in the development and axiomatization of the calculus of relations", *Studia Logica*, 50: 421-455.

Maddux, Roger, 2006, *Relation Algebras*, Studies in Logic and the Foundations of Mathematics, 150, Elsevier Science, Amsterdam.

Marciszewski, Witold, 1981, "Abstraction Operators", in Marciszewski, Witold, ed. *Dictionary of Logic as Applied in the Study of Language*, Springer, Dortrecht.

Mormann, Thomas, 1994, "A Representational Reconstruction of Carnap's Quasianalysis", *Philosophy of Science Association*, 1: 96-104

Mormann, Thomas, 2003, "Synthetic Geometry and *Aufbau*", in Bonk, Thomas, ed. *Language Truth, Knowledge. Contributions to the Philosophy of Rudolf Carnap*, Springer, Dordrecht: 45-65.

Mormann, Thomas, 2006, "Between Heidelberg and Marburg. The *Aufbau*'s neo-Kantian origins and the AP/CP-Divide", *Sapere Aude! Revista de Filosofía. México, F.Z.*, 1: 22-50.

Mormann, Thomas, 2009, "New Work for Carnap's Quasi-Analysis", *Journal of Philosophical Logic*, 38, 3: 249-282

Mormann, Thomas, 2016, "Carnap's *Aufbau* in the Weimar Context", in Damböck, Christian ed., *Influences on the* Aufbau, Vienna Circle Institute Yearbook, 18: 115-137

Mormann, Thomas, 2018, "Structure-preserving representations, constitution, and the relative a priori", *Synthese*, DOI: 10.1007/s11229-018-1759-y.

Mundy, Brent, 1986, "On the General Theory of Meaningful Representation", *Synthese*, 67: 391-437.

Proust, Joëlle, 1986, *Questions de forme: Logique et proposition analytique de Kant à Carnap*, Fayard, Paris ; Eng. tr. by Anastasios Brenner 1989, *Questions of form, logic and the analytic proposition from Kant to Carnap*. University of Minnesota Press, Minneapolis.

Quine, Willard Van Orman, 1969, *Ontological Relativity and Other Essays*, Open Court, LaSalle and Chicago.

Ramharter, Esther, Georg Rieckh, 2006, *Die Principia Mathematica auf den Punkt gebracht*, öbv und hpt, Wien.

Rheinberger, Hans-Jörg, 2010, *On historicizing epistemology. An essay*, Stanford University Press, Stanford.

Richardson, Alan, 1998, *Carnap's Construction of the World: The Aufbau and the Emergence of Logical Empiricism*, Cambridge University Press, Cambridge.

Russell, Bertrand, 1914, *Our Knowledge of External World as a Field for Scientific Method in Philosophy*, Open Court, London.

Schilpp, Paul, 1963, ed. *The Philosophy of Rudolf Carnap*, The Library of Living Philosophers, 11, Open Court, La Salle, Illinois.

Schnädelbach, Herbert, 1983, *Philosophie in Deutschland 1831–1933,* Suhrkamp Verlag, Frankfurt am Main; Eng. tr. by Eric Matthews 1984, *Philosophy in Germany 1931-1933*, Cambridge University Press, Cambridge.

Simmons, Harold, 2011, *An introduction to Category Theory*, Cambridge University Press, Cambridge.

Stubenberg, Leopold, 2016, "Neutral Monism", in Edward Zalta ed., *The Stanford Encyclopedia of Philosophy*, Winter 2016, <https://plato.stanford.edu/archives/win2016/entries/neutral-monism/>.

Suppes, Patrick, 2002, *Representation and invariance of scientific structures*, CSLI Publications, Stanford.

Swoyer, Chris, 1991, "Structural Representation and Surrogative Reasoning", *Synthese* 87: 449-508

Textor, Mark, 2021, "Mach's Neutral Monism", *HOPOS. The Journal of the International Society for the History of Philosophy of Science*, 11, 1: 143-165.

Tully, Robert, 2003, "Russell's Neutral Monism", in Griffin, Nicholas, ed. *The Cambridge Companion to Russell*, Cambridge University Press, New York: 332-371

van Fraassen, Bas, 2006, *Scientific Representation: Paradoxes of Perspective*, Clarendon press, Oxford.

Wille, Rudolf, Bernhard Ganter, 1999, *Formal Concept Analysis. Mathematical Foundations*, Springer-Verlag, Berlin.

Wishon, Donovan, 2021, "Radical Empiricism, Neutral Monism, and the Elements of Mind", in Strawson, Galen, ed. *Monism*, *The Monist*, 104, 1: 125-151, DOI: 10.1093/monist/onaa026.

Ziche, Paul, 2016, "Theories of order in Carnap's *Aufbau*", in Damböck, Christian ed., *Influences on the* Aufbau, Vienna Circle Institute Yearbook, 18: 77-98.

Ziehen, Theodor, 1913, *Erkenntnistheorie auf psychophysiologischer und physikalischer Grundlage*, Verlag von Gustav Fischer, Jena.

Ziehen, Theodor, 1920, *Lehrbuch der Logik*, A. Marcus & E. Webers Verlag, Bonn.

Quasi-analysis
A method to order non-homogeneus sets by means of the theory of relations

Rudolf Carnap

Abbreviations: Al[1] : Analysis into components,
Ct : Component,
QAl : Quasi-analysis,
QCt : Quasi-component,
SC : similarity circle.[2]

1. *The task of quasi-analysis*

The nature of individual objects (henceforth "elements") of any domain (henceforth "set" of elements) can be indicated by means of two different methods. The first method indicates for any individual element the characteristics that belong to it or the components (Ct) that it is composed of. We call this method analysis into components (Al). This name is also appropriate for those indications of characteristics that do not analyse the object but its concept, considered as the totality of the characteristics of the object. The second method indicates the relations that hold between elements. We call this method relational description. Although each of the two methods offers several variants, these are, however, more or less similar among themselves. At the same time, the methods are basically different. Indeed, following the first, one can make a statement about an individual element without taking into account other elements. Following the second, instead, every statement concerns only the relations of an element to one or more other elements. The two methods

[1] Carnap uses several abbreviations in the manuscript with the aim of rendering the text readable more easily. In the original manuscript, the abbreviations are chosen in accordance with the German words. In this translation, we have rendered the abbreviations consistent with the English words. This choice is motivated by the purpose of bestowing new original clarity, readability and elegance on the English version of the text.

[2] Our translation choice of "similarity circle" and "similarity neighbourhood" (6-7), standing for Carnap's original phrases "*Familienklasse*" and "*Verwandtenklasse*", is unliteral. This choice is justified by the purpose of keeping the language of *Quasizerlegung* on a homogeneous line of expression with that of the subsequent philosophical debate on quasi-analysis and similarity structure.

philinq X, 1-2022, pp. 255-272
ISSN (print) 2281-8618-ETS doi: 10.4454/philinq.v10i1.404

could be labelled, respectively, as indication of individual properties and indication of relational properties.

> Examples of Al: description of the set of conic sections through an account of the characteristics of the individual sections; description of a curve through its coordinate equation, *i.e.*, by giving the ordinate for each point on the abscissa; description of a physical state through the values of one (or more) state variable for every position; chemical description of a given substance through its composition of chemical elements; list of historical persons with a statement of the dates of birth and death for each of them.
>
> Examples of relational description: description of a geometrical figure which consists of points and straight lines through an indication of the relations of incidence; description of a curve through its natural equation, *i.e.*, through an indication of the position of each element of the curve relative to the preceding ones; description of a physical state through spatio-temporal differential equations, *i.e.*, through the relation between the value of a state variable in some spatio-temporal point and its values in the spatio-temporal neighbourhood; description of a group of persons by means of a genealogy, *i.e.*, by giving their kinship relations.

In opposition to Al, the relational description has the advantage that it does not overstep the given domain of objects. The elements of the set to be described are, indeed, not analysed into components (Ct), whose set is generally not included in the given one. The relational description is, as it were, an "immanent approach". On the other hand, the relational description has the drawback of being ponderous in the approach to the individual elements themselves. One cannot, indeed, make a statement about an element without reference to other elements, which are again characterised only through reference to other elements, and so forth.

Now, a method will be discussed here that allows a relational description to transform a given description in such a way that retains the properties of the immanent approach and assumes the form of the analysis. Thus, a single approach to the elements is possible. This transformation is called quasi-analysis (QAl).

The simplest version of QAl considered in the following discussion can be applied everywhere, even where it seems desirable to switch to more complicated versions. Quasi-analysis starts from a relational description based on a symmetric, reflexive and non-transitive relation. Let us call this initial relation

"S" and two elements *a* and *b*, such that *aSb*, "similar elements", "similarity pair" or "*S*-pair".

If a relational description is based on a relation *P* whose properties are different from the above-mentioned ones (symmetry, reflexivity, intransitivity) or on many relations, *P, Q, …*, then one must take as basic relation *S* certain transformations of these other relations. One can, if necessary, apply several of the following transformations one after the other.

1) *P* is transitive, in particular:
   a) transitive and symmetric. Here a degenerate case occurs: all the elements are similar one to the other (with the exception of the isolated ones, i.e., the elements that are similar only to themselves). Therefore, the set is homogeneous. There are no distinguishing properties and, consequently, no possibility of order.
   b) *P* is transitive but not symmetric. One applies the transformation 2), whereby transitivity is also removed.
2) *P* is not symmetric. One defines $S = P \dot{\cup} \breve{P}$ Def.
3) *P* is not reflexive. One defines $S = P \dot{\cup} I \restriction C'P$ Def.
4) There are two basic relations, *P* and *Q*.
   a) $P \subset Q$ holds. Moreover, *P* is symmetric, transitive and reflexive, *Q* is symmetric and reflexive. One defines $S = P|Q|P$ Def.
   b) $P \subset Q$ holds, but *P* and *Q* do not have the properties required in a). One constructs from them two new relations, *R* and *V*, which have these properties. In particular, one introduces symmetry and transitivity by applying the transformations 2) and 3) and transitivity by applying the transformation: $R = P_*$ Def. One defines: $S = R|V|R$ Def.
   c) $P \subset Q$ does not hold. One defines: $S = P \dot{\cup} Q$ Def.
5) There are more than two basic relations.
   a) One of these relations, *T*, is implied by all the others: $U \subset T, W \subset T, Z \subset T, \ldots$ One defines: $P = U \dot{\cup} W \dot{\cup} Z, \ldots$ Def, which yields: $P \subset T$. Then, one applies (4a) or (4b) on these two relations.
   b) The condition in (a) is not satisfied. One divides the relations into two classes and takes both unions of these classes as new relations: $P = T \dot{\cup} U \dot{\cup} \ldots$ Def., $Q = W \dot{\cup} Z \dot{\cup} \ldots$ Def. If possible, the division takes place most conveniently when the constructed relations *P* and *Q* satisfy the condition in 4a) or, if this is not possible, the condition in 4b).

Let these transformation rules be here given in short without justification. That they lead to the desired result, i.e., that the new constructed relations and the relation *S* possess the required properties (symmetry, reflexivity, intransitivity) can be easily seen. According to the peculiarities of the case, other transformations, which lead to the same desirable result, turn out to be often more appropriate.

Now, the problem of quasi-analysis can be formulated as follows. A set of elements is given and for every element the list of its similar elements. Find a description of this set that uses only these indications, but assigns to the elements quasi-components (QCts) or quasi-characteristics in such a way that every individual element can be described by itself, without reference to other elements, according to its QCts. This problem can be solved by assigning to the elements QCts in such a way that two elements have a QCts in common if and only if they are similar. To minimize arbitrariness, it is required for QAl to state about the elements nothing more than what is already contained in the given lists. This way, two elements are represented as identical if and only if they are identical according to the given lists. Finally, following the principle of economy, it is required that no unnecessary QCts occurs in QAl. For this reason, in order to find a QAl of a given set of elements, a method must satisfy the following four basic requirements [axioms, *translator's note*].

The four basic requirements.

(**I**) If two elements are similar, then they share at least one QCt.

(II) If two elements are not similar, then they do not share any QCt.

(III) If two elements *a* and *b* are "similarity equivalent" (i.e., *a* is similar to exactly the same elements that are similar to *b*), then they are "QCt-equivalent" (i.e., *a* and *b* possess exactly the same QCts).

The converses of (I), (II) and (III) do not need to be introduced as requirements. Rather, they follow from the above-mentioned basic requirements; cf. theorems (1), (2), (3) *infra*.

(IV) There is no QCs whose removal leaves the requirements (I), (II) and (III) still satisfied.

It is shown through the method to be discussed in the sequel that these four requirements are consistent one with the other and are all satisfied. It can be

seen that they are independent one of the other in analogy to the independence proof of the axioms of geometry (Hilbert).

> Every requirement is shown to be independent of the other by providing an example where it is not satisfied, but the others are. In the following example (I) is not satisfied, namely for *b* and *c*, while the other basic requirements are satisfied. *a, b, c, d* are elements and *aSb*, *bSc*, *cSd* hold (henceforth, we do not explicitly mention *aSa*, etc., and *bSa*, etc., since these similarities come respectively from the reflexivity and the symmetry of *S*). Also, α is a QCt of *a* and *b*, β is a QCt of *c* and *d.* In the next example, (II) is not satisfied, but the other requirements are satisfied. *a, b, c* are elements. *aSb*, *bSc* hold and α is a QCt of *a*, *b* and *c*. In the next example, only (III) is not satisfied, namely for any pair of elements. *a*, *b*, *c* are elements. *aSb*, *bSc*, *cSa* hold. Also, β and γ are the QCts of *a*, γ and α those of *b* and α and β those of *c.* In the next example, only (IV) is not satisfied, namely for β. *a, b, c, d* are elements. *aSb, bSc, cSa, cSd* hold. Also, α is QCt of *a*, *b* and *c*, β is QCt of *a* and *b*, γ of *c* and *d*.

The following seven theorems are consequences of the four basic requirements.

(1). Theorem. Let the four basic requirements be satisfied. If two elements share a common QCt, then they are similar.

This theorem is the converse of (I) and follows form (II).

(2). Theorem. Let the four basic requirements be satisfied. If two elements have no QCt in common, then they are not similar.

This theorem is the converse of (II) and follows from (I).

(3) Theorem. Let the four basic requirements be satisfied. If two elements are QCt-equivalent or similarity equivalent, then they are similar.

The second and the first parts of the theorem come from reflexivity of *S* and Theorem (1), respectively.

(4). Theorem. Let the four basic requirements be satisfied. If two elements are QCt-equivalent, then they are also similarity equivalent.

This theorem is the converse of (III).

> **Proof**. Let us assume that *a* and *b* are QCt-equivalent. By (I), for every element *c* which is similar to *a*, *a* and *c* share at least a certain QCt. Further, it follows from the QCs-equivalence of *a* and *b* that *b* and *c* share the same QCt. Hence, by (1), *bSc*. Therefore, every element which is similar to *a* is also similar to *b*.

(5). Theorem. Let the four basic requirements be satisfied. If two elements are not similarity equivalent, then they are not QCt-equivalent.

It follows from (4).

(6). Theorem. Let the four basic requirements be satisfied. If two elements are not QCt-equivalent, then they are not similarity equivalent.

It follows from (III).

(7). Theorem. Let the four basic requirements be satisfied. There is no QCt which is a "companion" of another QCt (a Ct or a QCt α is a companion of the Ct, or the QCt, β when α belongs only to elements to which β also belongs).

It comes from (IV), since such a QCt can be removed without violating (I) and (II).

2. *The first part of QAl: the similarity circles*

At first sight, the problem of QAl seems easy to solve: the relation of sharing a QCt between two elements easily takes the place of the given similarity relation. The difficulty lies in the fact that the similarity relation is not transitive while the relation of sharing a QCt is. Therefore, the attempt to finding a QAl by assigning a QCt α to an element *a* and the same QCt α to the elements *b, c, d* which are similar to *a* is clearly unsuccessful. It would indeed violate the requirement (II), since, from *aSb* and *aSc*, one cannot conclude *bSc*.

The subsequent discussion is intuitively based on a concrete **example**. The example is taken from the domain of the phenomenology of sense impression. During other researches, the investigation of this domain precisely suggested the development of the quasi-analytical method.

> **Example**. Let a set of 12 sounds, namely chords and individual tones, be given. Let us label the sounds, the elements of the set, *h, i, k, … t*: *h*= tone *d*, *i* =chord *d-f-a*, *k*= *c-e-g*, *l*= *c-e*, *m*= *f-a*, *n*= *d-f, o*= *c-e-a*, *p*= *c-f*, *q*= *c*, *r*= *d-a*, *s*= *g*, *t*= *c-g*. Let this composition be unknown. The 12 elements must be considered as indivisible and analysable only through quasi-analysis. For this, according to what we have said above, only the list of the pairs of elements which stand in the symmetric, reflexive, intransitive relation *S* are required. As such S-pairs let us introduce: *hi*, *hn*, *hr*, *im*, *in*, *io*, *ip*, *ir*, *kl*, *ko*, *kp*, *kq*, *ks*, *kt*, *lo*, *lp*, *lq*, *lt*, *mn*, *mo*, *mp*, *mr*, *np*, *nr*, *op*, *oq*, *or*, *ot*, *pq*, *pt*, *qt*, *st*. The opposite pairs *ih, nh, … ts,* do not need to be mentioned here, because of the symmetry of *S* (thus, we don't distinguish them from the former in the sequel). Similarly, we don't explicitly enumerate (reflexivity of *S*!) the identity *S*-pairs: *hh, ii, .. tt.* As one sees, *S* corresponds in the example to the familiar kinship of sounds, which is ordinarily called "agreement in (at least) a constituent tone" or "(at least) partial identity". This is emphasised here, however, only for the purpose of illustration, quasi-analysis does not pay attention to it. The elements must be indivisible in quasi-analysis which is based only upon the *S*-pairs. By using the example it can be shown that not only its results, but also many single steps of QAl are analogous to those of Al. This is another justification for choosing this name.

We call $\vec{S}'x$ the similarity neighbourhood[3] of x, i.e., the class of elements that are similar to x.

In our **example** we have: $\vec{S}'$h = [h, i, n, r], $\vec{S}'$i = [h, i, m, n, o, p, r], $\vec{S}'$k = [k, l, o, p, q, s, t], $\vec{S}'$l = [k, l, o, p, q, t], $\vec{S}'$m = [i, m, n, o, p, r], $\vec{S}'$n= [h, i, m, n, p, r], $\vec{S}'$o = [k, l, m, o, p, q, r, t], $\vec{S}'$p = [i, k, l, m, n, o, q, t], $\vec{S}'$q = [k, l, o, p, q, t], $\vec{S}'$r = [h, i, m, n, o], $\vec{S}'$s = [k, s, t], $\vec{S}'$t = [k, l, o, p, q, s, t].

(The following sections (8)-(12) and (16) are not required to apply the QAl method. Rather, they serve only to compare QAl with Al. Those who are interested only in the method and not in its justification can also omit them).

(8). Definition. $E_1 = \hat{x}\ \hat{y}\ (S'x \subset S'y)$ Def.

Thus, xE_1y means: the similarity neighbourhood of x is contained in the similarity neighbourhood of y.

Hence, we have in our **example**, $\vdash lE_1o$ but not $\vdash oE_1l$. Moreover, $\vdash sE_1t, \vdash qE_1t, \vdash hE_1i, \vdash mE_1i$.

Thus, a not symmetric relation E_1 is obtained from the symmetric relation S. This represents, in QAl, the following relation of Al: "y contains all the Cts of x".

(9). Definition. $E_2 = E_1 \dot{\cap} \breve{E_1}$ Def.

Thus, xE_2y means: E_1 and its converse hold between x and y. E_2 is the already mentioned relation of similarity equivalence. Indeed, we have:

(10). Theorem. $\vdash :. (x, y): xE_2y. \equiv . \vec{S}'x = \vec{S}'y$

According to the basic requirement (III), there arises the problem of finding a QAl where any two elements which stand in the relation E_2 are QCt-equivalent. To compare it again with the Al, E_2 correspond to the Ct-equivalence of Al in many cases. This, however, does not always occur, namely, not if a Ct occurs in Al only as companion, which, by (7), is excluded in QAl.

In the **example** $\vdash \vec{S}'l = \vec{S}'q$ hold. Thus, $\vdash lE_1q$ and $\vdash qE_1l$, whence $\vdash lE_2q$. Hence l and q must be QCt-equivalent. But they are not Ct-equivalent, since the tone e belongs, as Ct, to l, but not to q. We shall later on see that this disagreement is dependent upon (7), and thence upon the basic requirement (IV).

(11). Definition. $E = E_1 \dot{\cap} - \breve{E_1}$ Def.

We have that E if E_1, but its converse does not hold. The relation E of QAl represents the following relation of Al: "x is Ct-equivalent to a proper part of y".

[3] For this translation choice see fn 2

In the **example**, $\vdash l\ E\ o$, since $\vdash lE_1o.o - E_1l$. In Al, the Cts of *l* (the tones *c, e*) are proper parts of the Cts of *o* (*c, e, a*).

(12). Definition. $E_3 = \hat{x}\,\hat{y}\,\left(xE_1y\,.\,\vec{E}'x = \Lambda\right)$ Def.

Thus, xE_3y means: the similarity neighbourhood of *x* is contained in the similarity neighbourhood of *y* and there is no element which stands in the relation *E* to *x*. The relation E_3 of QAl represents the following relation of Al: "to *x* belongs only one Ct and the same belongs to *x*" or "*x* is Ct-equivalent to an individual Ct of *y*" or, in short, "*x* is an individual Ct of *y*".

In the **example** we had $\vdash sE_1t$ and $\vdash qE_1t$. However, there is no element *u* in the given set such that *u*E*s* or *u*E*q*. Hence, $\vdash qE_3t$ and $\vdash s\ E_3t$ hold. For comparison, let us consider again Al. We have: *q* (namely the tone *c*) and *s* (tone *g*) are individual Cts of *t*.

One may conjecture that one could carry out QAl by assigning to every element *x* (as it were the "totality of its individual QCts") the class of the elements which stands in the relation E_3 to it, i.e., the class $\overrightarrow{E_3}'x$. However, as an example can clearly show, such a method would result in a violation of the basic requirement (I).

In the **example** one had: $\vdash hE_1i$ and $\vdash mE_1i$. There is no element *u* of the set for which *u*E*h* or *u*E*m* hold. Thus, $\vdash hE_3i$ and $\vdash mE_3i$. There is no other element like *h* and *m* which stands in the relation E_3 to *i*: $\overrightarrow{E_3}'i = [h,m]$. On the other hand, neither *h* nor *m* belong to $\overrightarrow{E_3}'p$. Indeed, $\vdash h - S\ p$ and $\vdash m - E_1p$, thus $\vdash m - E_3p$. Therefore, if one take $\overrightarrow{E_3}'x$ as the class of the QCses of *x*, then the elements *i* and *p* would share no QCts. But since $\vdash iSp$ the basic requirement (I) would not be satisfied.

If we compare with the Al, then we clarify the reason for the shortcoming of a QAl based on E_3: not every Ct occurs somewhere in isolation, i.e., as the only Ct of an element.

The attempted QAl of *i* into the individual QCts *h* and *m* would correspond to an analysis of the chord *d-f-a* into the two Cts *d* and *f-a*. Indeed, *f-a* would not be further analysable here since the tones *f* and *a* never occur in isolation. Now, as we will see, in this case the QAl precisely represents the tripartition of Al.

We define the "class of similarity circles (SC)":

(13). Definition. $sim = \hat{\beta}\,(x,y\ \in\ \beta\ \supset\ xSy :.(z):(u).u \in \beta\ \supset\ uSz.\supset.z \in \beta)$ Def.

In words: we say β is a SC if the following conditions are satisfied: 1) any two members of β are similar; 2) if some *z* is similar to every member *u* of β,

then z belongs to β. A SC is thus a complete set of elements that are similar one to the other.

To a SC corresponds, in Al, a class that contains all and only those elements which share a given Ct. However, not all such classes of Al have a corresponding SC in QAl, namely, not if the class is based on a Ct which is a companion of another. One cannot provide any analogue to such a class if the basic requirement (IV) must not be violated.

In the **example**, the class [*k,l,o*] of elements, which share the tone *e* as Ct, is not a SC. For the latter condition of the definition is not satisfied: q does not belong to the set, but is similar to all its members. Since, later on, we draw the QCts out of the SCs, there is no QCt corresponding to the tone *e*. The reason for the above-considered disagreement between QAl and Al in relation to the tone *e* lies in the fact that *e* occurs only in elements in which the tone *c* also occurs. Therefore, *e* is a companion of *c*. Thence, according to the basic requirement (IV) and theorem (7), there is no QCt corresponding to it. On the other hand, $\delta = [h, i, n, r]$ is a SC. Its members are the elements that share the tone *d* as common Ct. Likewise, $\alpha = [i, m, o, r]$ is a SC, where the common Ct is the tone *a*. Similarly, $\gamma = [k, l, o, p, q, t]$ is a SC with the common tone *c*. One finds as remaining SCs: $\varphi = [i, m, n, p]$, $\xi = [k, s, t]$, $\pi = [i, m, o, p]$, $\rho = [i, m, n, r]$. Thus, $\vdash sim = [\delta, \alpha, \gamma, \varphi, \xi, \pi, \rho]$.

(14). Theorem. $\vdash :. (x, y): xSy. \supset . (\exists \alpha). \alpha \in sim. x, y \epsilon \alpha$

In words: for every pair of similar elements there is (at least) a SC which contains them both.

Proof. If there is no element *u* which is similar to *x* and *y* and different from both, then for $\alpha = [x, y]$ the theorem is satisfied. However, if there is such a *u*, then one constructs the class [*x,y,u*]. If there is no element *v* which is similar to every element of the class, then for [*x,y,u*] the theorem is satisfied. Otherwise, let us construct the class [*x,y,u,v*]. Continuing in this way, if the class of elements is finite, we must arrive, in a finite number of steps, at a class for which the theorem is satisfied.

(15). Theorem. $\vdash :: (x. y. \alpha): . \vec{S}'x = \vec{S}'y. \ \alpha \in sim. \supset : x \in \alpha. \equiv . y \in \alpha$

In words: if two elements are similarity equivalent, then one belongs to the same SCs of the other.

Proof. $\alpha \in sim$ and $x \in \alpha$ yield: any member of α is similar to x. Thus, since, $\vec{S}'x = \vec{S}'y$, it is also similar to y. Therefore, by (13), $y \in \alpha$.

(16). Definition. Repr = $\hat{x}\,\hat{\beta}(\beta \in sim :. (y): y \in \beta. \supset . xE_3y)$ Def.

X Repr β means: *x* is a member of β and stands in the relation E_3 to every other member of β. In this case, we call the element *x* a "representative" of the

SC β. The relation *Repr* of QAl corresponds in Al to the relation: "β contains all and only those members which share the only Ct of x". As in the above-indicated cases, exceptions to this correspondence arise when the Ct in question is a companion of another. There are also SCs without representatives. In Al this means: the common Ct of the set does not occur in isolation. Since $\mathsf{D}'\,Repr$ indicates the class of those SCs which have at least a representative, $sim - \mathsf{D}'Repr$ indicates therefore the class of the SCs without representatives. Also, there are SCs with many representatives. These representatives in Al are either Ct-equivalent or they differ from one another through Cts that are companions of other Cts.

In the **example**, for the SC $\mathfrak{d} = [h, i, n, r]$ we have $\vdash h\ Repr\ \delta$. Indeed, $\vdash \delta \in sim.h \in \delta.hE_3i.hE_3n.hE_3r$. In the Al: the Cs of *h*, namely the tone *d*, is what the members of $\mathfrak{d}$ have in common. Hence the expression "representative of δ" for *h*. A SC without representatives is, for example, $\alpha = [i,m,o,r]$, since none of its members stands in the relation E_3 to the other three. The reason for this is clarified by Al. The Cs common to all the members of α, namely the tone *a*, never occur in isolation. A SC with many representative is $\gamma = [k,l,o,p,q,t]$. Indeed, $\vdash l\ Repr\ \gamma.q\ Repr\ \gamma$ holds. In the Al, *l* and *q* differ from one another through the tone *e* as Ct, which is the companion of *c*.

as in the const. th.

Here again we could fall into a method that, at first sight, seems to be suitable for a QAl but does not lead to the desired result. Couldn't we assign to every element *x*, as QCts, the SCs to which it belongs? One clearly deduces from (13) that the essential basic requirements, namely (I), (II), (III) would be here satisfied. However, the attainment of the requirement of economy, (IV), would not be guaranteed. This can be very easily shown through the example. As we shall see, by avoiding this mistake one brings QAl and Al into a closer analogy.

In the **example**, according to the above-considered method, one would assign to the individual elements the following classes of SCs as their (attempted) classes of QCts: the class $[\delta]$ (i.e., the class whose only member is the SC δ) to the element *h*, the class $[\delta,\varphi,\alpha,\pi,\rho]$ to the element *i*, $[\gamma,\xi]$ to *k*, *l*: $[\gamma]$, *m*: $[\varphi,\alpha,\pi,\rho]$, *n*: $[\delta,\varphi,\rho]$, *o*: $[\gamma,\alpha,\pi]$, *p*: $[\gamma,\varphi,\pi]$, *q*: $[\gamma]$, *r*: $[\delta,\alpha,\rho]$, *s*: $[\xi]$, *t*: $[\gamma,\xi]$. The basic requirements (I), (II), (III) are here satisfied. But (IV) is not satisfied. For, even though we delete from the just given formulation the QCts π or ρ, or even all two, (I), (II), (III) still remain satisfied. These two SCs are also those without any analogue in the Al. While to the SCs δ, α, γ, φ, ξ correspond the classes of the elements sharing the tones *d, a, c, f, g*, respectively, neither the members of π nor those of ρ share a common Ct.

3. *The second part of QAl: the QCs-class as a result*

(17). Definition. $\mathrm{N} = \hat{\alpha}\,\widehat{\beta}\,(\alpha \in \beta.\beta \subset sim :: (\exists x, y): x \neq y.x, y \in \alpha:.(\gamma): \gamma \in \beta.x, y \in \gamma.\supset.\gamma = \alpha)$ Def.

Thus, $\alpha N \beta$ means: β is a class of SCs; the SC α belongs to β; there are two different elements in α such that it is not the case that both of them belong to another SC in β. In this case we say: "α is relatively necessary with respect to β". From this follows: $\vec{N}'\beta$ is the class of the relatively necessary SCs with respect to β. If β contains all the SCs (β=*sim*), then $\vec{N}'sim$ is the class of the "absolutely necessary SCs". $\alpha N sim$ means: α is an absolutely necessary SC.

> In the **example**, let us now take $\beta = [\gamma, \phi, \alpha, \pi]$. Thus, $\vdash \gamma N \beta$ holds, since the *S*-pair *kt* and others occur only in γ, but not in ϕ, α, π. Similarly, $\vdash \phi N \beta$ because of the pair *mn*; $\vdash \alpha N \beta$ because of the pair *ir*. However, $\vdash \pi - N\beta$. Therefore $\vdash \vec{N}'\beta = [\gamma, \phi, \alpha]$. Moreover, $\vdash \vec{N}'sim = [\gamma, \delta, \phi, \alpha, \xi]$ holds, i.e., these five SCs are absolutely necessary SCs, in particular γ because of the pair *kl*; δ because of *hi*; φ because of *np*; α because of *or*; ξ because of *ks*. On the other hand, π and ρ are not absolutely necessary SCs, for none of them contains a *S*-pair which does not belong also to other SCs.

(18). Theorem. $\vdash N \in \varepsilon$

It follows from (17). In words: whatever stands in the relation *N* to a class belongs to it as member.

(19). Theorem. $\vdash:.(\lambda, \kappa, \alpha): \alpha N \lambda.\kappa \subset \lambda.\supset.\alpha N \kappa$

In words: a SC is relatively necessary with respect to a class of SCs whenever it is relatively necessary with respect to another class of SCs where the former is contained (as subclass).

> Let us now take for **example** $\lambda = [\gamma, \phi, \alpha]$ and $\beta = [\gamma, \phi, \alpha, \pi]$, as before. Then, $\vdash.\gamma N \beta.\phi N \beta.\alpha N \beta$ yields $\vdash.\gamma N \lambda.\phi N \lambda.\alpha N \lambda$

(20). Theorem. $\vdash.\vec{N} \mid \vec{N} = \vec{N}$

In words: the class μ of the relatively necessary SCs with respect to another class is identical with the class of the SCs that are relatively necessary with respect to μ itself.

(21). Theorem. $\vdash.\vec{N}'(\vec{N}'sim) = \vec{N}'sim$

It follows from (20). In words: the class $\vec{N}'sim$ of the absolutely necessary SCs is identical with the class of the SCs that are relatively necessary with respect to $\vec{N}'sim$ itself.

(22). Definition. $\mathrm{suf} = \hat{\beta}(\beta \subset sim:.(x, y): x \neq y.xSy.\supset.(\exists\alpha).\alpha \in \beta.x, y \in \alpha)$ Def.

In words: β is said to be a "sufficient class of SCs" if for any two similar elements there is a SC α in β that contains them both.

(23). Theorem. $\vdash :. (\alpha, \beta): \alpha \in suf . \alpha \subset \beta . \supset . \beta \in suf$

In words: a class is a sufficient class of SCs whenever it contains a sufficient class of SCs as its subclass.

In the **example**, let us take $\kappa = [\gamma, \delta, \phi, \alpha, \xi]$. Hence, $\vdash \kappa \in suf$. Indeed, for any two similar elements there is at least one of these five SCs that contains them both. Let us take $\kappa_1 = [\gamma, \delta, \phi, \alpha, \xi, \pi]$, $\kappa_2 = [\gamma, \delta, \phi, \alpha, \xi, \rho]$, $\kappa_3 = [\gamma, \delta, \phi, \alpha, \xi, \pi, \rho]$ (=sim). By (23) we have $\vdash \kappa_1, \kappa_2, \kappa_3 \in hinr$. In our example there is no other sufficient class of SCs: $\vdash suf = [\kappa, \kappa_1, \kappa_2, \kappa_3]$.

(24). Theorem. $\vdash :. (x, y, \beta): xSy . \beta \in suf . \supset . (\exists \alpha) . \alpha \in \beta . x, y \in \alpha$

It follows from (22). In words: for any two similar elements there is (at least) a SC in every sufficient class of SCs that contains them both.

(25). Theorem. $\vdash \vec{N}'sim \subset p'suf$

Follows from (17) and (22). In words: the absolutely necessary SCs belong to the intersection of sufficient classes of SCs and therefore to each sufficient class of SCs.

In the **example** we have $\vdash p'suf = \kappa$, since $\vdash . \kappa \subset \kappa_1 . \kappa \subset \kappa_2 . k \subset \kappa_3$. Further, $\vdash \vec{N}'sim = \kappa$ implies $\vdash \vec{N}'sim \subset \kappa$.

(26). Definition. $nec = \hat{\beta}(\beta \subset \vec{N}'\beta)$ Def.

In words: β is said to be a "relatively necessary class of SCs" ($\beta \in nec$) if it contains only those classes that are relatively necessary with respect to β itself.

In the **example** we had $\vdash . \gamma N\lambda . \phi N\lambda . \alpha N\lambda$, for $\lambda = [\gamma, \phi, \alpha]$. Therefore $\vdash \lambda \in nec$ holds, i.e., λ is a relatively necessary class of SCs. Let us take $\lambda_1 = [\alpha, \xi, \pi, \rho]$, so we have again $\vdash \lambda_1 \in nec$. Indeed, $\vdash \alpha N\lambda_1$ because of the pair *or*, which only occurs in α, but not in ξ, π or ρ. Similarly, $\vdash \xi N\lambda_1$ because of the pair *st*; $\vdash \pi N\lambda_1$ because of *ip*; $\vdash \rho N\lambda_1$ because of *in*. For the above indicated classes $\kappa, \kappa_1, \kappa_2, \kappa_3$, $\vdash \kappa \in nec$ holds. Indeed, we have $\vdash \gamma N\kappa$, since the S-pair *kl* occurs only in γ, but not in δ, φ, α, ξ. Moreover, $\vdash \delta N\kappa$ because of the *S*-pair *bn*; $\vdash \phi N\kappa$ because of *np*; $\vdash \alpha N\kappa$ because of *or*; $\vdash \xi N\kappa$ because of *st*. On the other hand, $\vdash \kappa_1 \sim \in nec$, for $\pi - N\kappa_1$, i.e., there is no S-pair in π which does not occur also in (at least) one of the remaining classes γ, δ, φ, α, ξ. Similarly, $\vdash \kappa_2 \sim \in nec$, for $\rho - N\kappa_2$. In the same way, $\vdash \kappa_3 \sim \in nec$.

(27). Theorem. $\vdash :. (x): x \in nec . \supset . \vec{N}'x = x$

Follows from (26) and (28). In words: a relatively necessary class of SCs is identical with the class of the relatively necessary classes with respect to itself.

(28) Theorem. $\vdash \vec{N}'sim \in nec$

It follows from (21). In words: the class of the absolutely necessary SCs is a relatively necessary class of SCs.

(29). Definition. *sn* = *suf* ∩ *nec* Def.

If $\beta \in sn$, then β is said to be a "QAl-class". Thus, *sn* is the class of the QAl-classes. Such a QAl-class contains exactly sufficiently many SCs, to apply a QAl satisfying the four basic requirements, and no unnecessary SC.

(30). Theorem. $\vdash . \ sn \subset suf$

In words: every QAl-class is a sufficient class of SCs.

(31). Theorem. $\vdash . \ sn \subset nec$

In words: every QAl-class is a relatively necessary class of SCs. Both theorems follow from (29).

(32). Theorem. $\vdash . sn \subset \mathrm{Cl}'sim$

Follows from (29) and (22). In words: every QAl-class is a class of SCs.

(33). Theorem. $\vdash : . (x, y, \beta) : xSy . \beta \in sn . \supset . (\exists \alpha) . \alpha \in \beta . x, y \in \alpha$

Follows from (30) and (24). In words: for any two similar elements there is (at least) a SC in every QAl-class that contains them both.

(34). Theorem. $\vdash . \vec{N}'sim \ \subset p'sn$

It follows from (25). In words: the absolutely necessary SCs belong to every QAl-class. In general, *sn* has many members, i.e., many QAl-classes, so that there are many possible QAl of a given field of relations. An example of it will be discussed in the sequel (with the rules for a practical application of QAl). The following three propositions are concerned with the case where only one QAl is possible.

(35). Theorem. $\vdash : \vec{N}'sim \in suf . \supset . \vec{N}'sim \in sn$

Follows from (28).

(36). Theorem. $\vdash : \vec{N}'sim \in suf . \supset . sn \in 1$

By (34), every member of *sn* contains $\vec{N}'sim$. Now if *sn* contains more than one member, then the member of *sn* which is different from $\vec{N}'sim$ (for example ξ) must contain both $\vec{N}'sim$ and at least a SC which is not contained in $\vec{N}'sim$. Since, by assumption, any two similar elements already occur in at least one member of $\vec{N}'sim$, ξ would not belong to *nec* and therefore, by (31), not to *sn* either.

(37). Theorem. $\vdash : \vec{N}'sim \in suf . \supset . \vec{N}'sim = (\imath x)(x \in sn)$

Follows from (35) and (36). The content of the propositions (35), (36), (37) is: if the class of the absolutely necessary SCs is a sufficient class of SCs, then 1) it is itself a QAl-class, so 2) there is only one QAl-class and hence, in this case, 3) this class is the only QAl-class.

The premise of this proposition is fulfilled in our **example**. Indeed, we had: $\vdash \vec{N}'sim = [\gamma, \delta, \phi, \alpha, \xi]$ and for $\kappa = [\gamma, \delta, \phi, \alpha, \xi] \vdash \kappa \in suf$.Therefore, κ is the

only QAl-class. This can be concluded from the previous results also: $\vdash . suf = [\kappa, \kappa_1, \kappa_2, \kappa_3]$, $\vdash \kappa \in nec$, $\vdash \kappa_1, \kappa_2, \kappa_3 \sim\in nec$ yield, by (29), $\vdash . sn = [\kappa]$.

If there are many QAl-classes, then we must decide which of them forms the basis of the QAl. If one of them has a smaller number of element than the others, then we will choose it. Otherwise, the choice is of no consequence, if it is not determined for reasons outside QAl, which are here not under discussion.

The QAl on the basis of a QAl-class assigns to every element the members of this class (i.e., SCs) as its QCts. Hence, the class of the QCts assigned to an element is called its "QCt-class". These QCt-classes for the individual elements are the result of the quasi-analysis. Actually, one can speak of the QCt-class of an element only as far as a definite QAl-class is concerned, i.e., a definite member of *sn*, since *sn* generally has multiple members. Let <*β*> indicate the relation of a QCt-class to its element with respect to the QAl-class β, then we must define:

(38). Definition. $< \beta > = \hat{x}\, \hat{y}(\beta \in sn.\, x \subset \beta.\, y \in p'x$ Def

Thus, ⊐'<*β*> indicates the class of the QCs-classes with respect to the QAl-class *β*.

(39). Theorem. $\vdash : (\beta, x).< \beta >' x \subset \beta$

It follows from (38). In words.: the QCt-class of an element with respect to a QAl-class *β* is contained in *β*.

In the **example** we had κ as the only QAl-class. Therefore, we assign the SCs (namely, the members of κ) to their respective members as QCts: accordingly, the QCt γ to the elements *k, l, o, p, q*, t; the QCt *δ* to the elements *h, i, n, r*; φ to the elements *i, m, n, p*; α to the elements *i, m, o, r*; *ξ* to the elements *k, s, t*. Thus, the following QCt-classes (with respect to *κ*) are assigned to the following individual elements: <*κ*>'*h* = [*δ*], i.e., δ is the only QCt of *h*; *h* is a representative of the QCt *δ*; <*κ*>'*i* = [*δ*, *φ*, *α*], we say: the element *i* "consists", or "is composed of", the QCts *δ*, *φ*, *α*. Moreover, <*κ*>'*k* = [*γ*, *ξ*]; <*κ*>'*l* = [*γ*]; <*κ*>'*m*=[*φ*, *α*]; <*κ*>'*n*= [*δ*, *φ*]; <*κ*>'*o*= [*γ*, *α*]; <*κ*>'*p*=[*γ*, *φ*]; <*κ*>'*q*=[*γ*] <*κ*>'*r*=[*δ*, *α*]; <*κ*>'*s*=[*ξ*]; <*κ*>'*t*=[*γ*, *ξ*].

4. *Testing the procedure against the four basic requirements*

(40). Theorem. $\vdash :. (x, \alpha, \beta): x \in \alpha. \alpha \in \beta. \beta \in sn. \equiv . \alpha \in < \beta >' x$

Follows from (38). In words: if an element belongs to a member of a QAl-class, then this member belongs to the QCt-class of the element with respect to the same QAl-class; and vice versa. This gives rise to the following propositions.

(40a). Theorem. $\vdash :. (x, \alpha, \beta) : \alpha \in < \beta >' x . \supset . x \in \alpha$

(40b). Theorem. $\vdash :. (x, \alpha, \beta) : \alpha \in < \beta >' x . \supset . \alpha \in \beta$

(40c). Theorem. $\vdash :. (x, \alpha, \beta) : \alpha \in < \beta >' x . \supset . \beta \in sn$

(40d). Theorem. $\vdash :. (x, \alpha, \beta) : \alpha \in < \beta >' x . \supset . \alpha \in sim$

Follows from (40c), (40b), (32).

(41). Theorem. $\vdash :. (x, y, \beta) : xSy . \beta \in sn . \supset . (\exists z) . z \in < \beta >' x . z \in < \beta >' y$

Proof. Since $\beta \in sn$, we have $\beta \in suf$ by (29). It follows from (22) that for any two similar elements, in particular for the elements x and y of the premise, there is a SC α in β which contains them both. Thus, for $z = \alpha$, we have $z \in < \beta >' x . z \in < \beta >' y$, by (40).

(42). Theorem. $\vdash :. (x, y, \beta) : (\exists z) . z \in < \beta >' x . z \in < \beta >' y . \supset . xSy$

Proof. By assumption and (40a), we have $x \in z$ and $y \in z$. By (40d), we also have $z \in sim$. Thus, *xSy* follows from (13).

(43). Theorem. $\vdash :. (x, y, \beta) : xSy . \beta \in sn . \equiv . (\exists z) . z \in < \beta >' x . z \in < \beta >' y$

Proof. One of the two conditionals, out of which this equivalence is composed, corresponds to (41). In the other direction, i.e., the converse of (41), (42) yields *xSy* and (40c) yields $\beta \in sn$.

(44). Theorem. $\vdash :. (x, y, \beta) : \vec{S}'x = \vec{S}'y . \beta \in sn . \supset . < \beta >' x = < \beta >' y$

(Indirect) **proof**. Suppose the proposition is false. Then, there is a member of the QCt-class (with respect to β) of an element, say *x*, which does not belong to the QCt-class of *y*. Let α be this member. Thus, we have $\alpha \in < \beta >' x . \alpha\sim \in < \beta >' y . \alpha \in \beta$. It follows from (40a) that $x \in \alpha$ and $\alpha \in sim$ follows from (40d). $\alpha \in \beta, \beta \in sn$ and (40) yield $y\sim \in \alpha$, which contradicts $\vec{S}'x = \vec{S}'y$, by (15).

(45). Theorem. $\vdash :. (x, y, \beta) : < \beta >' x = < \beta >' y . \supset . \vec{S}'x = \vec{S}'y . \beta \in sn$

Proof. By assumption and (3), we have: $\beta \in sn$. Hence, by (41), $(u) : uSx . \supset . (\exists v) . v \in < \beta >' u . v \in < \beta >' x$. Now, by assumption again, $(z) . z \in < \beta >' x . \supset . z \in < \beta >' y$. Therefore, we have $(u) : uSx . \supset . (\exists v) . v \in < \beta >' u . v \in < \beta >' x . v \in < \beta >' y$. Since, by applying (42), we obtain $(\exists v) . v \in < \beta >' u . v \in < \beta >' y . \supset . uSy$, we have $(u) . uSx \supset uSy$, thus $\vec{S}'x \subset \vec{S}'y$. Similarly, we prove $\vec{S}'y \subset \vec{S}'x$. Therefore, $\vec{S}'x = \vec{S}'y$.

(44) e (45) can be summarized in the following theorem:

(46). Theorem. $\vdash :. (x, y, \beta) : \vec{S}'x = \vec{S}'y . \beta \in sn . \equiv . < \beta >' x = < \beta >' y$

(47). Theorem. $\vdash :: (\alpha, \beta) :. \alpha \in \beta . \beta \in sn . \supset : (\gamma) : (\exists x, y) . xSy . x \neq y . x, y \in \gamma . \gamma \in \beta . \supset . \gamma = \alpha$

In words: if α belongs to the QAl-class β, then there are two similar elements which belong only to α and not to any other member of β different from α.

Proof. By assumption and (31), we obtain $\beta \in nec$. Thus, by (27), $\beta = \vec{N}'\beta$, whence $\alpha \in \vec{N}'\beta$ and therefore $\alpha N \beta$. By applying (17), the theorem holds.

The theorems derived above lead to the result that the basic requirements (I)-(IV) are satisfied. By (41), (I) is satisfied.

If, following (39), we examine the above indicated QCt-classes of the elements of our **example**, then we can confirm that they share at least one QCt for the S-pairs mentioned at the beginning of the example. Indeed, we have: the QCt-classes of *c* and *h* share δ, those of *h* and *r* share δ, those of *h* and *r* share δ, those of *l* and *m* α and φ, and so forth.

That the basic requirement (II) is satisfied follows from (42), using the contrapositive of its conditional (i.e., exchanging and negating both of its parts). In summary, (43) means that (I) and (II) are satisfied.

The examination of the example shows that no other pairs of elements than *S*-pairs share a QCt.

(44) means that the basic requirement (III) is satisfied. The converse of (III), which is also required, but does not need, as was shown above, to be indicated by itself, is satisfied by (45). That (III) and its converse are satisfied is summarized in (46).

In our **example**, only the pairs *kt* and *lq* are similarity equivalent. The same pairs are also QCt-equivalent: $< \kappa >' k =< \kappa >' t = [\gamma, \xi], < \kappa >' l =< \kappa >' q = [\gamma]$.

(47) means that for any QAl-class, and therefore any possible QAl, and any QCt α there is a *S*-pair whose elements have no other QCt in common than α. It follows that, whatever QCt one removes, there would be a *S*-pair whose elements would have no QCt to share. Hence, the basic requirement (I) would not be satisfied. Therefore, the basic requirement (IV) is satisfied.

In our **example**, in order to prove $\kappa \in nec$, following (26), we have shown that for any member of κ there is a *S*-pair occurring only in this and not in the other members. Therefore, whatever member of κ we remove, there would be two similar elements with no QCt to share anymore. This would violate (I). Therefore, in our example, the requirement (IV) is satisfied.

5. *Comparison of quasi-analysis with analysis*

The analogy between both methods has been repeatedly emphasised at each step. Using the example, it can be very easily shown that the result of QAl, in

opposition to the simpler possibilities mentioned earlier, not only satisfies the four basic requirements, but it also achieves a more detailed correspondence to Al. It turns out that all the Cts have a QCt as their strict analogue, with the exception of the Cts that are companions of others and are not allowed, therefore, to have any analogue in a QAl that satisfies the four basic requirements.

In our **example**, among the Cts of the Al, namely the tones *c, d, e, f, a,* the tone *e* is a companion of *c*. Thus, it has no corresponding QCt. With this latter exception, the QAl corresponds strictly to the Al. If, in the QCt-classes of the elements that the QAl has found, we replace the QCt γ by the tone *c*, similarly δ by *d*, φ by *f*, ξ by *g*, α by *a*, then we have $< \kappa >' h = [\delta]: [d]$, $< \kappa >' i = [\delta, \phi, \alpha]: [d, f, a], < \kappa >' k = [\gamma, \xi]: [c, g]$, and so forth. Accordingly, for every element the result of quasi-analysis is the chord or the individual tone that the analysis assigns to it.

Edizioni ETS
Palazzo Roncioni - Lungarno Mediceo, 16, I-56127 Pisa
info@edizioniets.com - www.edizioniets.com
Finito di stampare nel mese di febbraio 2022