

PHENOMENOLOGY AND MIND





# PHENOMENOLOGY AND MIND

## METHODS OF PHILOSOPHY

*Edited by Stefano Bacin and Francesca Boccuni*



*Phenomenology and Mind* practices double blind refereeing and publishes in English.

#### SCIENTIFIC BOARD

##### **Phenomenology and Social Ontology (PERSONA)**

Lynne Baker †

Stefano Besoli (Università di Bologna)

Jocelyn Benoist (Université de Paris 1- Sorbonne)

Daniele Bruzzone (Università Cattolica Sacro Cuore, Piacenza)

Amedeo G. Conte (Università degli studi di Pavia, Accademia Nazionale dei Lincei)

Paolo Costa (Fondazione Bruno Kessler, Trento)

Vincenzo Costa (Università degli studi del Molise)

Guido Cusinato (Università degli studi di Verona, Max Scheler Gesellschaft)

Paolo Di Lucia (Università degli studi di Milano)

Giuseppe Di Salvatore (Fondazione Camprostrini, Verona)

Maurizio Ferraris (Università degli studi di Torino)

Elio Franzini (Università degli studi di Milano)

Vanna Iori (Università Cattolica Sacro Cuore, Piacenza)

Shaun Gallagher (University of Memphis, University of Central Florida; Københavns Universitet;

University of Hertfordshire)

Vittorio Gallese (Università degli studi di Parma)

Dieter Lohmar (Universität zu Köln)

Giuseppe Lorini (Università degli studi di Cagliari)

Verena Mayer (Ludwig-Maximilians-Universität München)

Lorenzo Passerini Glazel (Università degli studi di Milano-Bicocca)

Jean-Luc Petit (Université de Strasbourg, Laboratoire de Physiologie de la Perception et de l'Action, Collège de France, Paris)

Stefano Rodotà †

Paolo Spinicci (Università degli studi di Milano)

Corrado Sinigaglia (Università degli studi di Milano)

Massimiliano Tarozzi (Università degli studi di Trento)

Dan Zahavi (Institut for Medier, Erkendelse og Formidling, Københavns Universitet)

Wojciech Żelaniec (Uniwersytet Gdański)

##### **Ethics and Political Theory (CeSEP)**

Giampaolo Azzoni (Università degli studi di Pavia)

Elvio Baccarini (University of Rijeka)

Carla Bagnoli (Università degli studi di Modena e Reggio Emilia)

Gaia Barazzetti (Université de Lausanne)

Francesco Battegazzorre (Università degli studi di Pavia)

Antonella Besussi (Università degli studi di Milano)

Alessandro Blasimme (INSERM UMR1027 - Université Paul Sabatier, Toulouse)

Alberto Bondolfi (Fondazione Bruno Kessler, Trento)

Patrizia Borsellino (Università degli studi di Milano-Bicocca)

Francesco Botturi (Università Cattolica di Milano)

---

© The Author(s) 2018.

La presente opera, salvo specifica indicazione contraria, è rilasciata nei termini della licenza Creative Commons Attribution 4.0 Unported (CC BY 4.0: <https://creativecommons.org/licenses/by/4.0/legalcode>).

CC 2018 Firenze University Press

Università degli Studi di Firenze

Firenze University Press

Borgo Albizi, 28, 50122 Firenze, Italy

[www.fupress.com](http://www.fupress.com)

*Phenomenology and Mind*, on-line: <http://www.fupress.net/index.php/pam>

Stefano Canali (Scuola Internazionale Superiore di Studi Avanzati - SISSA)  
Ian Carter (Università degli studi di Pavia)  
Emanuela Ceva (Università degli studi di Pavia)  
Antonio Da Re (Università degli studi di Padova)  
Mario De Caro (Università degli studi di Roma III)  
Corrado Del Bò (Università degli studi di Milano)  
Emilio D'Orazio (POLITEIA - Centro per la ricerca e la formazione in politica ed etica, Milano)  
Maurizio Ferrera (Università degli studi Milano)  
Luca Fonnesu (Università degli studi di Pavia)  
Anna Elisabetta Galeotti (Università del Piemonte Orientale, Vercelli)  
Barbara Herman (University of California, Los Angeles - UCLA)  
John Horton (Keele University)  
Andrea Lavazza (Centro Universitario Internazionale di Arezzo)  
Eugenio Lecaldano (Università degli studi di Roma "La Sapienza")  
Neil Levy (University of Melbourne)  
Beatrice Magni (Università degli studi di Milano)  
Filippo Magni (Università degli studi di Pavia)  
Massimo Marassi (Università Cattolica di Milano)  
Alberto Martinelli (Università degli studi di Milano)  
Susan Mendus (University of York)  
Glyn Morgan (Syracuse University in New York)  
Anna Ogliari (Università Vita-Salute San Raffaele)  
Valeria Ottonelli (Università degli studi di Genova)  
Federico Gustavo Pizzetti (Università degli studi di Milano)  
Mario Ricciardi (Università degli studi di Milano)  
Nicola Riva (Università degli studi di Milano)  
Adina Roskies (Dartmouth College)  
Giuseppe Sartori (Università degli studi di Padova)  
Karsten R. Stueber (College of the Holy Cross)  
Nadia Urbinati (Columbia University)  
Corrado Viafora (Università degli studi di Padova)

#### **Cognitive Neurosciences, Philosophy of Mind and Language, Logic (CRESA)**

Edoardo Boncinelli (Università Vita-Salute San Raffaele)  
Stefano Cappa (Institute for Advanced Study, IUSS, Pavia)  
Benedetto de Martino (University College London, UCL)  
Claudio de' Sperati (Università Vita-Salute San Raffaele)  
Michele Di Francesco (Institute for Advanced Study, IUSS, Pavia)  
Massimo Egidi (Libera Università Internazionale degli Studi Sociali Guido Carli di Roma, LUISS Guido Carli, Roma)  
Francesco Guala (Università degli studi di Milano)  
Vittorio Girotto (Istituto Universitario di Architettura di Venezia, IUAV, Venezia)  
Niccolò Guicciardini (Università degli studi di Milano)  
Diego Marconi (Università degli studi di Torino)  
Gianvito Martino (Università Vita-Salute San Raffaele)  
Cristina Meini (Università del Piemonte Orientale)  
Martin Monti (University of California, Los Angeles, UCLA)  
Andrea Moro (Institute for Advanced Study, IUSS, Pavia)  
Michael Pauen (Berlin School of Mind and Brain, Humboldt-Universität)  
Massimo Piattelli Palmarini (University of Arizona)  
Giacomo Rizzolatti (Università degli studi di Parma)  
Marco Santambrogio (Università degli studi di Parma)  
Achille Varzi (Columbia University)  
Nicla Vassallo (Università degli studi di Genova)

## **History of Ideas, Aesthetics, Metaphysics (CRISI)**

*Massimo Adinolfi (Università degli studi di Cassino)*

*Simonetta Bassi (Università degli studi di Pisa)*

*Giovanni Bonacina (Università degli studi di Urbino)*

*Adone Brandalise (Università degli studi di Padova)*

*Enrico Cerasi (Università Vita-Salute San Raffaele)*

*Fabrizio Desideri (Università degli studi di Firenze)*

*Giulio D'Onofrio (Università degli studi di Salerno)*

*Roberto Esposito (Istituto Italiano di Scienze Umane-SUM, Napoli)*

*Adriano Fabris (Università degli studi di Pisa)*

*Romano Gasparotti (Accademia delle Belle Arti, Brera-Milano)*

*Sebastano Ghisu (Università degli studi di Sassari)*

*Dario Giugliano (Accademia delle Belle Arti, Napoli)*

*Giacomo Marramao (Università degli studi di Roma Tre)*

*Maurizio Migliori (Università degli studi di Macerata)*

*Salvatore Natoli (Università degli studi di Milano-Bicocca)*

*Pier Aldo Rovatti (Università degli studi di Trieste)*

*Vesa Oittinen (Università di Helsinki)*

*Giangiorgio Pasqualotto (Università degli studi di Padova)*

*Mario Perniola (Università degli studi Roma Tor Vergata)*

*Hans Bernard Schmid (Universität Basel)*

*Emidio Spinelli (Università degli studi La Sapienza-Roma)*

*Pirmin Stekeler-Weithofer (Universität Leipzig)*

*Italo Testa (Università degli studi di Parma)*

*Francesco Tomatis (Università degli studi di Salerno)*

*Federico Vercellone (Università degli studi di Torino)*

*Vincenzo Vitiello (Università Vita-Salute San Raffaele)*

*Frieder Otto Wolf (Freie Universität Berlin)*

*Günter Zöllner (Ludwig-Maximilians-Universität München)*



# CONTENTS

<i>Stefano Bacin, Francesca Boccuni</i> Introduction: Methods of Philosophy	10
--	----

## SECTION 1. ANALYTIC PHILOSOPHY

<i>Timothy Williamson</i> Model-Building as a Philosophical Method	16
---	----

<i>Francesca Boccuni</i> Reference in Formal Semantics and Natural Language: A Methodological Route	24
--	----

<i>Bianca Cepollaro</i> Intuizioni linguistiche e filosofia sperimentale: metodi a confronto	36
---	----

## SECTION 2. ETHICS

<i>Roger Crisp</i> The Methods of Ethics	48
---	----

<i>Massimo Reichlin</i> On the Idea of a 'Method' in Moral Philosophy	60
--	----

## SECTION 3. PHENOMENOLOGY

<i>Dermot Moran</i> What is the Phenomenological Approach? Revisiting Intentional Explication	72
--	----

<i>Sara Heinämaa</i> Two Ways of Understanding Persons: A Husserlian Distinction	92
---	----

<i>Roberta Lanfredini</i> Phenomenological Empiricism	104
--	-----

<i>Roberta De Monticelli</i> The Paradox of Axiology. A Phenomenological Approach to Value Theory	116
--	-----



# CONTENTS

## SECTION 4. PHILOSOPHY AND GENDER ISSUES

*Marina Sbisà*  
Ways to Be Concerned with Gender in Philosophy 132

*Laura Caponetto*  
Filosofia del linguaggio femminista, atti linguistici e riduzione al silenzio 146

## SECTION 5. HISTORY OF IDEAS

*Ethel Menezes Rocha*  
God, Eternal Truths and the Rationality of the World in Descartes 162

*Marcos André Gleizer*  
Spinoza on metaphysical doubt and the “Cartesian circle” 176

*Alfredo Gatto*  
Descartes e il problema della teodicea nella prima modernità 194

*Raffaele Ariano*  
Geografia e filosofia. Riflessioni su *Pensiero vivente* di Roberto Esposito a partire da Spinoza, Cavell e Foucault 212

*Corrado Claverini*  
Dove va la filosofia italiana? Riflessioni sull’*Italian Thought* 222

# INTRODUCTION

# INTRODUCTION

*Stefano Bacin, Francesca Bocconi*  
Introduction: Methods of Philosophy

---

STEFANO BACIN

*Università degli Studi di Milano*  
*stefano.bacin@unimi.it*

FRANCESCA BOCCUNI

*Università Vita-Salute San Raffaele*  
*boccuni.francesca@univr.it*

---

# INTRODUCTION: METHODS OF PHILOSOPHY

---

A critical examination of its method(s) belongs to the core business of philosophy from its beginnings. In recent discussion, however, metaphilosophical issues have gained an especially prominent role. As a constant flood of publications shows, metaphilosophy and philosophical methodology are among the areas in which the current debates are most lively.<sup>1</sup> The international conference hosted in 2017 by the Faculty of Philosophy of the Università Vita-Salute San Raffaele, and organised by Elisabetta Sacchi, Sarah Songhorian and Stefano Bacin with the support of the research centres CESEP, CRESA, CRISI, DIAPOREIN, GENDER, and PERSONA, aimed at gathering contributions on methodological issues in different fields of current philosophical studies: analytic philosophy, ethics, phenomenology, philosophy and gender issues, and history of ideas. The sections of this issue include most of the papers presented at that conference.

If there is a philosophical perspective that started as an intellectual endeavour based on a revolutionary methodology, which through decades brought original philosophical questions to light, it is analytic philosophy. In this spirit, Timothy Williamson scrutinises the analytic attitude towards testing philosophical theories via a falsificationist methodology. Though still widespread in the analytic community, a falsificationist methodology might be less fruitful and more prone to error than a more flexible alternative brought forward by Williamson, namely the reliance on model-building, carried over into the philosophical debate from empirical and social sciences.

Williamson's contribution is followed by two papers devoted to two methodological divides in the analytic tradition. The first divide, which is a more classical one, concerns whether, in the analysis of core philosophical notions such as the notion of reference of singular terms, a methodological approach in terms of formal theories and their models versus an approach focusing on the analysis of natural language is preferable. This divide informs Francesca Boccuni's contribution.

Bianca Cepollaro's article, on the other hand, concerns the linguistic intuitions connected with the use of certain expressives such as derogatory terms in the light of the second, more recent,

---

<sup>1</sup> See e.g. the very diverse picture provided by two recent handbooks: Cappelen, Szabó Gendler & Hawthorne (2016) and D'Oro & Overgaard (2017). For a more recent, significant sample of the debate, see Baz 2018, along with the review by Cappelen & Deutsch (2018), and Williamson (2007, 2018).

divide in the analytic tradition: whether the analysis of central philosophical notions should be carried out via a methodology relying exclusively on theoretical investigation (the so-called “armchair philosophy”) or via an experimental approach.

The title of this special issue echoes that of a philosophical masterpiece that put methodological issues in the foreground, thereby marking a perspective that is still very much at the centre of the debate, namely Henry Sidgwick’s *The Methods of Ethics*. Both papers of the section on ethics follow this connection and take their lead from Sidgwick’s views.

Roger Crisp examines Sidgwick’s understanding of ‘method’ and his methodological practice. After considering Sidgwick’s conditions of proper self-evidence of intuitions and some limits in his assessment of ‘dogmatic intuitionism’, Crisp suggests in this light a more constructive debate between advocates of different normative theories as a methodological guideline in moral philosophy.

Massimo Reichlin points out an issue in Sidgwick’s talk of a method of ethics that still affects current debates. Reichlin argues that Sidgwick’s construal of the three methods of ethics yields an infelicitous confusion between methods and normative theories of morality, which has in turn brought a consequentialist bias in the debates of the twentieth and twenty-first centuries.

In the first paper devoted to aspects of the phenomenological approach, Dermot Moran underscores the holistic nature of phenomenology. Moran examines the main features of phenomenology, characterising it as a philosophical method that, in opposition to objectivist naturalism, grounds on an investigation of the a priori relations between subjectivity and the different domains of objectivity.

Sara Heinämaa investigates a Husserlian distinction between two ways of understanding other persons, their motivations and actions: the empirical way and the genuinely intuitive way, crucially grounded, according to Heinämaa’s interpretation of Husserl, on self-understanding. Both self-understanding and understanding of others enable ethical reflections, self-shaping and judgements on others, by being not confined to the actual but by opening to the possibilities of change and renewal.

Roberta Lanfredini compares phenomenological empiricism with classical empiricism, on the one hand, and logical empiricism, on the other. She argues that phenomenology is a kind of radical, though not reductionist, empiricism that is to be assimilated neither to the first strain of empiricism nor to the second.

Roberta De Monticelli tackles some traditional philosophical issues revolving around the notion of value. She proposes a phenomenological answer in terms of Scheler’s material axiology to the vexed questions concerning the cultural versus the universal source of values and the truth and falsehood of value judgements. This proposal is spelled out in terms of a clarification of the bottom-up phenomenological approach, which De Monticelli supports via experimental phenomenology.

The fourth section of the volume is devoted to the debate concerning some philosophical considerations regarding gender issues, which have now established not merely as an object of philosophical analysis, but as a topic of metaphilosophical significance. This is fittingly illustrated in Marina Sbisà’s paper, which investigates several ways in which the discourse on gender can be tackled philosophically, examines the difficulties and limitations of each of them and proposes a consideration of gendered subjectivity in intersubjective relations understood as a basic process in human life.

Laura Caponetto focuses on the role that feminist philosophy of language has ascribed to Austin’s theory of speech acts to account for the notion of silencing in the debate on

pornography and censorship. The article's aim is to discuss Austin's theory in order to unveil a form of discourse injustice that has been largely undernoticed.

The final section gathers papers that do not directly address methodological issues, but exemplify a historical-critical approach. The paper included in this section revolve around two foci: central aspects of Descartes' epistemology and their reception in Spinoza, and the current debate on the so-called "Italian Theory".

In the first group of papers, Marcos Gleizer defends Spinoza's solution of the circle between evidence and that would affect Descartes' epistemology at its core. Against most readings, Gleizer argues that Spinoza's thesis of the "self-manifestation of reason" allows him to avoid the circularity between the truth of clear and distinct ideas, and God's existence.

Descartes' conception of the eternal truths is under scrutiny in the two following papers. Ethel Rocha considers some epistemological consequences of Descartes' conception of the divine creation of the eternal truths. Finally, Rocha argues that view seems to commit Descartes to the impossibility for the human mind to fully grasp the truth, but this in turn stands in tension with Descartes' entire metaphysical project. In his paper, Alfredo Gatto explores the connection between Descartes' thesis of God's free creation of the eternal truths and the idea of a theodicy, to argue that Leibniz' and Malebranche's project of a theodicy entails the rejection of the *basi tenets* of Descartes' thesis.

Finally, two papers are devoted to recent debates on Italian philosophy. Corrado Claverini sketches a genealogy of the recent debates on what has been labelled "Italian Thought" and the grounds of the growing international reception of Italian philosophy. Focusing on one defining moment in that debate, that is, Roberto Esposito's *Living Thought*, Raffaele Ariano points out in the reconstruction of an Italian philosophical tradition the striving to overcome the boundaries of academic philosophy, and the role played by the interpretation of Spinoza's thought in the ontological transformation of Michel Foucault's idea of biopolitics.

### REFERENCES

- Baz, A. (2018). *The Crisis of Method in Contemporary Analytic Philosophy*. Oxford: Oxford University Press;
- Cappelen, H., & Deutsch, M. (2018). Review of *The Crisis of Method in Contemporary Analytic Philosophy*, by Avner Baz, April. <https://ndpr.nd.edu/news/the-crisis-of-method-in-contemporary-analytic-philosophy/>;
- Cappelen, H., Szabó Gendler, T., & Hawthorne, J. (Eds.) (2016). *The Oxford Handbook of Philosophical Methodology*. Oxford, New York: Oxford University Press;
- D'Oro, G., & Overgaard, S. (Eds.) (2017). *The Cambridge Companion to Philosophical Methodology*. Cambridge: Cambridge University Press;
- Williamson, T. (2007). *The Philosophy of Philosophy*, Oxford: Blackwell Publishing;
- Williamson, T. (2018). *Doing Philosophy: From Common Curiosity to Logical Reasoning*. Oxford: Oxford University Press.

# SECTION

# 1



# SECTION 1

ANALYTIC PHILOSOPHY

*Timothy Williamson*

Model-Building as a Philosophical Method

*Francesca Boccuni*

Reference in Formal Semantics and Natural Language: A Methodological Route

*Bianca Cepollaro*

Intuizioni linguistiche e filosofia sperimentale: metodi a confronto

---

TIMOTHY WILLIAMSON

University of Oxford

[timothy.williamson@philosophy.ox.ac.uk](mailto:timothy.williamson@philosophy.ox.ac.uk)

---

# MODEL-BUILDING AS A PHILOSOPHICAL METHOD\*

---

## *abstract*

*The method of building simplified formal models of phenomena under study is widespread in contemporary natural and social science; much scientific progress consists in the provision of better models. A model-building methodology has also been used with success in analytic philosophy, for example by Carnap in his development of intensional semantics. Arguably, philosophers have overlooked how much progress their discipline has made through their failure to conceive it in model-building terms. By using the method more extensively, they can overcome the fragility to error inherent in the naïve falsificationist methodology on which many analytic philosophers rely.*

---

## *keywords*

*models, idealization, progress, Carnap, intensional semantics, falsificationism, Popper, error-fragility*

---

\* The ideas in this paper were presented at the 2018 conference on *The Methods of Philosophy* at Cesano Maderno; I thank participants for valuable discussion. Elsewhere I put this account of model-building in philosophy in the context of a general account of philosophical method (Williamson, 2018) and develop it further (Williamson, 2017).

### 1. Models in natural science

Many natural scientists aim at a distinctive kind of progress which philosophers are just starting to recognize as an appropriate aim for them too.

The stereotype of scientific progress is discovering a new law of nature. Such laws are meant to be universal generalizations about the natural world, holding without exception for all times and places, by some sort of necessity: nice, if you can find one. However, most natural science studies messy complex systems – cells, animals, planets, galaxies – which are hard to characterize by universal laws. What laws must hold of all tigers, for example? ‘All tigers are striped’ won’t do, because there are albino tigers. ‘All tigers are four-legged’ won’t do either, because there are three-legged tigers, and so on. ‘All tigers are animals’ is true, but doesn’t get us far. Although tigers obey the fundamental laws of physics, like everything else in nature, that won’t console a biologist who wants to say something specific about living things as contrasted with elementary particles and stars. If we keep watering down our initial attempts, we may eventually reach something exceptionless, but the danger is that it will be too weak and uninformative to be of much interest. This isn’t just a problem about animals. Complex systems of all shapes and sizes tend to be messy and unruly.

To manage the problem, scientists have revised their objectives. Instead of seeking universal laws about complex systems, they build simplified *models* of them. Occasionally these are physical models: water running through a sand tray to model a river eroding its banks, a construction of colored rods and balls to model a DNA molecule. More typically, the models are abstract, defined by mathematical equations which describe how a hypothetical system changes over time. The hypothetical system is vastly simpler than the real-life systems of interest, but still has a few of their key features. The strategy is to analyse the behavior of the hypothetical system mathematically, in the hope that it will simulate some puzzling aspects of the real-life systems’ behavior, and thereby cast light on them (Weisberg (2013) provides a good introduction to the philosophy of scientific model-building).

For example, you might wonder why a population of predators – say, foxes – and a population of prey – say, rabbits – keep oscillating, though rises and falls in one do not coincide with rises and falls in the other. A key point is that, holding other things equal, the more foxes there are, the more rabbits get eaten, but the more rabbits there are, the more fox cubs survive. One can write down differential equations that express the rate of increase or decrease of each population in terms of the current number of predators and prey. They are known as the Lotka-Volterra model. In most ways, it is grossly over-simplified: it ignores changes in the vegetation the rabbits feed on, changes in the tendency of humans to hunt foxes and

rabbits, variations among foxes, variations among rabbits, and so on. Since such factors make a difference, the equations are not universal laws. Indeed, they couldn't be, since for mathematical reasons the change in population is treated as continuous, even though in real-life it changes in whole numbers: when one of 200 rabbits dies, the number goes straight down to 199, with no intermediate time when the number of living rabbits was 199.5. Nevertheless, despite all these over-simplifications, the model correctly predicts some general structural features of population change in predator-prey species. Much progress in natural science is now of this kind. Once we have a successful model, we can try building a little more real-life complexity back into it, step by step, but the models will always be vastly simpler than real life itself – otherwise they would be too complex to analyze.

Sometimes there is no workable alternative to model-building. For example, biologists wonder why two-sex reproduction is the norm for animals, since reproduction by three sexes or none is possible in principle. If you want to understand why a phenomenon *doesn't* occur, you can't go out and observe and measure it. Instead, a good strategy is to build hypothetical models of the phenomenon to see what goes 'wrong' with it. You might study a model where both two-sex reproduction and three-sex or no-sex reproduction occur, to see which does better, perhaps in achieving genetic variation within the species, which enables it to adapt evolutionarily to changes in the environment. Such models aim not to predict observed quantities but to explain an absence.

Humans are a classic example of messy complex systems. In one way or another, much – though not all – of philosophy is about humans. Thus moral and political philosophy mainly concerns a good human life and a good human society. Philosophy of science concerns human science; philosophy of concerns human art; philosophy of language concerns human language. Though philosophy of art mind pays some attention to non-human animal minds, its main focus is on human minds, and in any case non-human animals are messy complex systems too. Though in principle epistemology concerns all knowledge, in practice it mainly concerns human knowledge. Logic and metaphysics are partial exceptions, since they tend to proceed at a level so fundamental that informative, precise, exceptionless laws are obtainable. For the rest, however, one might expect a model-building strategy to be appropriate.

That isn't how most philosophers have seen it. Many still aim at exceptionless laws, even about messy complex systems – humans – for whom natural scientists have mainly abandoned that ambition. In that respect, philosophers have done their field a disservice, by inadvertently setting it up for failure. People who contrast progress in natural science with deadlock in philosophy often do so on the basis of a false image of natural science. Failing to appreciate how much scientific progress consists in building better models, they fail to ask how much philosophical progress consists in building better models too.

One example of progressive model-building in philosophy is epistemic logic, which advances in just that way. Its models are not universal laws; they involve grossly unrealistic simplifications. Nevertheless, they cast light on human knowledge in the manner of a scientific model.

When philosophers work with probabilities, they typically use models without emphasizing the fact. For instance, a simple model of uncertainty is a lottery. To make things definite, suppose that exactly 100,000 tickets have been sold, numbered in order; there is just one winner, chosen at random. Thus if you have one ticket, its probability of losing is  $99,999/100,000$ . Which statements about the lottery should you accept? You might decide that requiring 100% certainty is unreasonably demanding, and resolve to accept just statements with a probability of at least 95%. Immediately, there is a problem. By your rule, you accept the statement that the winning number is at least 5,001 (because its probability is 95%), and you accept the statement that the winning number is at most 95,000 (because its probability

## 2. Models in philosophy

is 95%), but you refuse to accept the statement that the winning number is between 5,001 and 95,000 inclusive (because its probability is only 90%). Thus you accept each of two statements separately, but you refuse to accept the result of putting them together, their conjunction. A politician who did something like that on television in an election campaign could expect to get crucified. You might think that 95% was a bad choice of threshold for acceptance, and choose a different threshold. But a little calculation shows that the only thresholds for acceptance which avoid such problems, even when there are more tickets, are 0% and 100%. Since a threshold of 0% means accepting every statement whatsoever – total credulity – you are back to a threshold of 100%, the standard of certainty you already rejected as unreasonably demanding. Thus even such a ‘toy’ model can illustrate the difficulties of basing acceptance and rejection on information about probabilities.

If you think about the lottery model, you can quickly identify some of its simplifying assumptions. For instance, it assumes that you know exactly how many tickets have been sold. In practice, the organization running the lottery may not announce or even know how many tickets have been sold; even if they announce a number, you may give a nonzero probability to the hypothesis that they are mistaken or lying. Then you may also give a nonzero probability to the winning number being 100,001 (since more than 100,000 tickets may have been sold), and you may give a higher probability to the winning number being 1 than to its being 100,000 (since fewer than 100,000 tickets may have been sold). But taking account of all those realistic complications is not time well spent. Thinking about the simple model takes one more quickly to the heart of the problem. When more complex probabilistic models are needed to understand more intricate problems, mathematically-minded epistemologists construct them too.

In philosophy of language, model-building as understood here goes back at least to Rudolf Carnap. An example is his theory of the meaning of *modal operators*, words like ‘possibly’ and ‘necessarily’. He treated them as building up more complex sentences from simpler ones: thus from the sentence ‘Everything changes’ they make sentences such as ‘Possibly everything changes’ and ‘Necessarily everything changes’.

Logicians had already designed precise formal languages with symbols for logical words such as ‘not’, ‘or’, ‘and’, ‘something’, and ‘everything’, which enable one to build up more complex expressions from simpler ones, without limit. We can think of them as taking expressions as input and delivering more complex expressions as output. For example, if you input the sentence ‘Everything changes’ to ‘not’, the output is the sentence ‘Not everything changes’. Logicians also had a framework for analyzing the meaning of such general words. To each expression, simple or complex, they assigned something called its *extension*, encoding its application to the world. For instance, since the word ‘red’ applies to red things and not to non-red ones, its extension includes the former and excludes the latter. If a sentence applies to the world, its extension is truth; if it doesn’t apply to the world, its extension is falsity. ‘Not’ makes an output sentence the opposite in extension of the input sentence: if ‘Everything changes’ is true then ‘Not everything changes’ is false, while if ‘Everything changes’ is false then ‘Not everything changes’ is true. In effect, the extension of ‘not’ just swaps truth and falsity. ‘Or’, ‘and’, ‘something’, and ‘everything’ work similarly: they each transform the extension of the input into the extension of the output. Operators which do that are called *extensional*. For each expression of the language, such rules determine its extension from the extensions of the simple words from which it is built up. That helps explain how we can understand complex sentences we never previously encountered by understanding the familiar words of which they are made and how they are put together. In effect, such extensional semantics is an elementary – but very powerful – model of linguistic meaning, though people did not think of it like that at the time.

Carnap wanted to add symbols for ‘possibly’ and ‘necessarily’ to the formal language. His problem was that such modal operators don’t fit the previous model of meaning: they are not extensional.

For suppose that ‘possibly’ is extensional. Then if I pick a sentence ‘X’ and don’t tell you what it is, but only whether it’s true, you should be able to work out whether ‘Possibly X’ is true. In one case you can do that: if I tell you that ‘X’ is true, you can work out that ‘Possibly X’ is true too, since actuality implies possibility. But if I tell you that ‘X’ is *false*, you can’t work out whether ‘Possibly X’ is true. I haven’t given you enough information; the answer depends on what ‘X’ is. For example, if ‘X’ is ‘Napoleon won at Waterloo’ (false), then ‘Possibly X’ is true: although Napoleon lost, he could have won. But if ‘X’ is ‘5 is more than 6’ (also false), then ‘Possibly X’ is false too: 5 could not have been more than 6. Thus the extension of ‘X’ doesn’t always determine the extension of ‘Possibly X’. ‘Possibly’ is not extensional.

Similarly, suppose that ‘necessarily’ is extensional. Then you should be able to work out whether ‘Necessarily X’ is true. In one case you can do that: if I tell that ‘X’ is false, you can work out that ‘Necessarily X’ is false too, since necessity implies actuality. But if I tell you that ‘X’ is *true*, you can’t work out whether ‘Necessarily X’ is true. I haven’t given you enough information; the answer depends on what ‘X’ is. For example, if ‘X’ is ‘Napoleon lost at Waterloo’ (true), then ‘Necessarily X’ is false. But if ‘X’ is ‘6 is more than 5’ (also true), then ‘Necessarily X’ is true too. Thus the extension of ‘X’ doesn’t always determine the extension of ‘Necessarily X’. ‘Necessarily’ is not extensional.

Carnap solved the problem by considering not just extensions in the *actual* world, the way things are, but profiles of extensions over all *possible* worlds, ways things could have been. He borrowed the idea of possible worlds from Leibniz, though he preferred to use more linguistic entities, ‘state-descriptions’. He called the profiles *intensions*. For example, since the extension of ‘Napoleon lost at Waterloo’ is truth in every world in which Napoleon lost at Waterloo and falsity in every world in which Napoleon did not lose at Waterloo, the intension of ‘Napoleon lost at Waterloo’ assigns truth to each of the former worlds and falsity to each of the latter ones. Carnap’s crucial insight was that although the extension of the input to a modal operator doesn’t always determine the extension of the output, the *intension* of the input *does* always determine the intension of the output. He gave rules for calculating the latter in terms of the former. He interpreted ‘possibly’ as ‘in some possible world’ and ‘necessarily’ as ‘in every possible world’.

In more detail, Carnap’s rule for ‘possibly’ is that if the input is true in some possible world, then the output is true in *every* possible world, while if the input is false in every possible world, then the output is also false in every possible world. Thus ‘Napoleon won at Waterloo’ is true in some possible world, so ‘Possibly Napoleon won at Waterloo’ is true in every possible world. But ‘5 is more than 6’ is false in every possible world, so ‘Possibly 5 is more than 6’ is also false in every possible world. Thus the intension of the input determines the intension of the output; ‘possibly’ is intensional rather than extensional.

For ‘necessarily’, the rule is that if ‘X’ is true in every possible world, then ‘Necessarily X’ is also true in every possible world, whereas if ‘X’ is false in some possible world, then ‘Necessarily X’ is false in *every* possible world. Thus ‘Napoleon lost at Waterloo’ is false in some possible world, so ‘Necessarily Napoleon lost at Waterloo’ is false in every possible world. By contrast, ‘6 is more than 5’ is true in every possible world, so ‘Necessarily 6 is more than 5’ is also true in every possible world. Thus the intension of the input determines the intension of the output; ‘necessarily’ is intensional rather than extensional.

Since the rules for extensional operators like ‘not’, ‘or’, ‘and’, ‘something’, and ‘everything’ work for extensions in any possible world, Carnap easily adapted them to calculating intensions. The upshot was a complete intensional semantics for his whole formal language: every formula,

however complex, has an intension, determined step by step from the intensions of the simple constituents out of which it is composed. It's a significantly more sophisticated model of meaning than extensional semantics. Through the work of Richard Montague, David Lewis, and many others, Carnap's intensional semantics has massively influenced both philosophy of language and semantics as a branch of linguistics. Although the models have become ever more elaborate, they preserve the crucial move from extensions to intensions.

Carnap worked in a more model-building spirit than his predecessors. He didn't construct his formal language to do mathematics in, or to reveal the hidden essence of all languages. He constructed a simple model language to demonstrate a way for modal operators to work. As we learn ever more of the extraordinary complexity underlying even the most ordinary conversations, philosophers of language and linguists may have to rely increasingly on a model-building methodology.

### **3. Working models, counterexamples, and error-fragility**

Models are fun. You can play with them. That's not just an incidental side benefit; it's what they are for, in both natural science and philosophy. We learn by manipulation, playing about: if you can't manipulate the real thing, a good second-best is often to manipulate a model of it. You can fiddle with this or that component, changing it slightly to see what difference it makes, what varies with what. That way you come to understand more deeply how the model works. If the model is any good, you thereby come to understand better how the real thing works too. For instance, you can't arbitrarily change how English works, to see what difference it makes, but you can arbitrarily change the rules of an artificial language, and calculate the consequences.

To be easily manipulated, a model should be defined in mathematically or logically precise and tractable terms. If the definition is vague, or too complicated, its consequences are unclear: one has to fall back on one's prior philosophical instincts to guess how it behaves, instead of using the model to test those instincts. By contrast, a well-defined model allows one to calculate rigorously how it and variations on it behave, bypassing those prior instincts, and so to learn something unexpected. With a model-building methodology, rigor and playfulness go naturally together.

The rigor of model-building is not the rigor most philosophers are used to. Traditional philosophical rigor requires dismissing a claim once a counterexample to it has been given. In that sense, most models are born refuted, because they involve false simplifying assumptions. For instance, models in epistemic logic typically treat agents as logically perfect. Some philosophers dismiss those models accordingly.

In physics, models of the solar system may treat a planet as a point mass, as if all its mass were concentrated at its center. Of course, physicists know quite well that planets are not point masses and do not behave exactly like them. Nevertheless, physicists do not dismiss such models, for they also know that much can be learned from them. By contrast, if one tried to write into the model a fully accurate description of the planet, with all its craters and bumps, the result would be too complicated to permit calculation. It takes skill to distinguish amongst the features of a model those which have lessons to teach us from those which are mere artefacts of the need to keep things simple. Philosophers are having to learn that skill.

To many philosophers, dismissing the true counterexample rather than the false generalization seems like a disregard for truth. It would indeed be intellectually irresponsible to go on *believing* the generalization in the face of a clear counterexample. But that's not the model-building attitude. One can recognize that a generalization is both false and a key component of a model that points us towards genuine truths.

If counterexamples don't refute a model, what does? Within the model-building methodology, what displaces a model is a better model. Part of its superiority may be that it

deals more adequately with counterexamples to the old model, but it should also reproduce in its own way the old model's successes. A new model with that combination of virtues may be very hard to find.

Model-building contrasts with the methodology of *conjectures and refutations*, championed by Karl Popper. On the crude version of his view, scientists put forward bold conjectures, informative universal generalizations, which can be falsified but can never be verified. A single negative instance, a counterexample, will falsify the generalization; no finite number of positive instances will verify it. Scientists do their utmost to refute it, by finding such a counterexample. Once it is refuted, they put forward another bold conjecture, and so on. One problem for such a falsificationist methodology, in both natural science and philosophy, is that it is *error-fragile*. In other words, a single mistake can have disastrous consequences. For suppose that we are testing a bold conjecture, and take ourselves to have found a counterexample. As good falsificationists, we dismiss the conjecture and go on to the next one. But what if the counterexample was a mistake? We are fallible; sometimes we misjudge single instances. In that case, the original conjecture may have been true after all. But we never return to it; we are too busy testing new bold conjectures. Philosophers' reliance on counterexamples can be alarmingly close to crude falsificationism: once a counterexample is accepted, there's no going back on it. By contrast, the model-building methodology is much less error-fragile, for it gives no such decisive power to a single judgment. Models are compared over a variety of dimensions.

None of this means that philosophy should go over entirely to a model-building methodology. In some areas, such as logic, we have found many true and informative universal generalizations. In others, good models may be too much to expect. Even where good models are available, as in epistemology, we may do best by using *both* methodologies. For if each independently pulls in the same direction, that's stronger evidence that it's the right direction. Such a combination of methodologies is more robust, unless they pull in opposite directions. The potential of the model-building methodology for philosophy is only beginning to be explored. Its scope and limits should be clearer fifty years from now.

#### REFERENCES

- Carnap, R. (1947). *Meaning and Necessity*. Chicago: University of Chicago Press;
- Weisberg, M. (2013). *Simulation and Similarity: Using Models to Understand the World*. Oxford: Oxford University Press;
- Williamson, T. (2017). Model-building in philosophy. In R. Blackford and D. Broderick (Eds.) *Philosophy's Future: The Problem of Philosophical Progress* (pp. 159-173). Oxford: Wiley Blackwell;
- Williamson, T. (2018). *Doing Philosophy: From Common Curiosity to Logical Reasoning*. Oxford: Oxford University Press.





---

FRANCESCA BOCCUNI

*Università Vita-Salute San Raffaele*

*boccuni.francesca@univr.it*

---

# REFERENCE IN FORMAL SEMANTICS AND NATURAL LANGUAGE: A METHODOLOGICAL ROUTE

---

## *abstract*

*In this paper, I will tackle the notion of reference of singular terms in the light of a classic analytic divide, i.e. whether its analysis, like the analysis of other basic notions, should be carried out in natural language or in the semantics of formal frameworks. I will incline toward the latter strategy, and consider reference in classical first-order logic as the simplest framework in which to investigate reference.*

---

## *keywords*

*reference, arbitrariness, formal languages, natural language, formal semantics*

- 1. Overview** The history of analytic philosophy is scattered with issues related to fundamental notions, which are hard to settle and, at the same time, such that (some) scholars do not seem to agree upon the most sensible methodology to tackle them.
- By *methodology*, I mean the choice between the two horns of a very classic divide in the analytic tradition, i.e. whether in order to tackle those issues it is more reasonable to investigate them within and by *formal languages* or within and by *natural language*. In this respect, there are two radically opposing views: by the first view, one might translate natural language into a formal language, in order to unveil the significance of those basic notions, whose analysis in natural language can be subject to the ambiguities typical of this latter kind of context;<sup>1</sup> by the second, formal languages are supposed to just model notions whose significance and relevance cannot be appreciated but in natural language.
- The aim of this paper is to tackle the notion of *reference* of singular terms, in particular, with respect to whether its analysis, like the analysis of other basic notions, should be carried out in natural language or in the semantics of formal frameworks. I will incline toward the latter strategy, and consider reference in classical first-order logic as the simplest framework we have at disposal.
- In §2, I will provide two classic examples of fundamental notions with respect to the proposed methodological divide. In §3, I will then focus on the notion of reference of singular terms, in order to provide some basis for the main claim that the semantics of formal languages, in particular of classical *first-order logic*, is the right source for the analysis of reference—in particular, supporting a deflationary view of reference, i.e. *arbitrary reference* (§4). In order to carry out this project, a few more issues will be addressed: what is meant by *arbitrariness* and whether arbitrary reference is genuine reference (§5); what semantics for the case study analysed, i.e. classical first-order logic, can be provided (§6). I will then conclude by mentioning some further issues that will need to be tackled in future research and by taking stock of the main claims in this paper (§7).

---

<sup>1</sup> This strategy reminds one of Frege's renown scorn for the use of natural language in the investigation of the foundations of mathematics. This attitude brought him to formulate his *Begriffsschrift* in 1879 and, consequently, his colossal and, alas!, fatally flawed *opus magnum*, i.e. the *Grundgesetze der Arithmetik*, 1893-1903.

Let us consider two examples to clarify what is at stake: the notion of *truth*, and the notion of *quantification* in formal languages—in particular with respect to the debate on the ontological commitment of classical second-order logic.

The notion of truth is one of those fundamental notions analytic philosophers have investigated at length, and for the analysis of which different strategies have been envisaged. With respect to the proposed divide, truth may be investigated by recurring to formal tools,<sup>2</sup> regardless of whether this analysis is somehow faithful to the apparent use we make of it in natural language. Consider, for instance, the work of Alfred Tarski, who proved in 1936 the inexpressibility of truth in classical first-order logic. Tarski showed that, as soon as we add a truth-predicate to first-order logic in order to express the truth and falsity of sentences of that language *in that very language*, the so-called *Liar paradox* arises.<sup>3</sup> Tarski's solution to the paradox was to construe a hierarchy of more and more expressive formal languages, each one, apart from the lowest, containing the truth predicate for the language immediately weaker. Tarski's reliance on formal resources implied a rather pessimistic consideration of natural language: natural language is deemed as intrinsically inconsistent, since it is semantically closed, i.e. it contains its own truth-predicate and, thus, reproduces the Liar paradox.<sup>4</sup> With respect to the divide I am proposing, if one agrees with Tarski on natural language, then the correct notion of truth should be the one regimented in a, possibly axiomatic, formal theory of the appropriate form. A second strategy to deal with the Liar paradox might be to revise one or more principles of the underlying logic – whether formal or informal, in order to comply with the manifest functioning of truth in natural language. This latter strategy is the one championed by, e.g., Kripke (1975) and Priest (2006). Though their respective solutions to the Liar paradox are indeed very different logically,<sup>5</sup> they both agree that Tarski's solution provides a rather unnatural reading of the behaviour of the notion of truth in natural language.<sup>6</sup> On the face of the possible strategies on the market, then, it seems reasonable to ask: where does *the* notion of truth come from originally?<sup>7</sup> And by what resources, whether formal or informal, are we supposed to provide an analysis of it?

A further, though somehow restricted, example comes from the philosophy of logic, in particular from the debate on second-order quantification and its ontological commitment. Second-order logic contains two sorts of variables, first- and second-order, and related quantificational resources. The former variables and quantifiers vary over a domain of individuals, the latter vary over a domain of second-order entities – properties, Fregean concepts, sets, classes, or the like. As Quine (1986) famously complained, second-order logic is set-theory in disguise, since its second-order variables are taken to vary over a domain of sets of first-order individuals.<sup>8</sup> To be fair, Quine's claim, in order to be motivated, hinges on two presupposed Quinean slogans, namely

## 2. The Methodological Divide: Truth and Quantification

---

2 This approach is underwritten by the many authors who have been working on axiomatic formal theories of truth, from possibly different logical strains, e.g. Hartry Field, Volker Halbach, Andrea Cantini, Leon Horsten, just to mention a few.

3 Famously, the sentence stating its own falsity, which informally might be rendered as e.g. (L) The sentence L is false.

4 And many other paradoxes, for that matter.

5 Kripke (1975) relies on a strong Kleene three-valued logic; Priest (2006), on a three-valued paraconsistent logic.

6 Namely a distinction in infinitely many languages, which, altogether, should capture what truth is. See for instance Kripke (1975, p. 695): “Unfortunately this picture seems unfaithful to the facts. If someone makes such an utterance as (1) [Most (i.e., the majority) of Nixon's assertions about Watergate are false.], he does not attach a subscript, explicit or implicit, to his utterance of 'false', which determines the 'level of language' on which he speaks.”

7 Supposing it would be a great advantage, as I tend to think it would, if there were only one.

8 The correct quotation is “set theory in sheep's clothing”. See Quine (1986, ch. 5).

- (1) to be is to be the value of a variable;
- (2) since intensional entities like properties or Fregean concepts are “creatures of darkness”,<sup>9</sup> whereas sets have a clear identity criterion in the principle of extensionality, sets are to be preferred over intensional entities as the values of second-order variables in modelling second-order logic.

Both presuppositions are questionable to me, still, setting my qualms aside, a received view provides a set-theoretic semantics for second-order logic, by this, committing second-order logic to the existence of mathematical entities. So, for any sentence of natural language that can be expressed in second-order logic, the logical form and ontological commitment of that fragment of natural language are ultimately unveiled by the appropriate formal resources. Boolos (1984, 1985) championed an alternative view on second-order logic, aimed at vindicating its ontological innocence. He provided a semantics in terms of *plural quantification*, whose naturalness and primitivity is justified by examples taken from natural language,<sup>10</sup> (fragments of which are) used as a paradigmatic model for formal languages. To the best of my knowledge, Boolos did not explicitly claim that natural language is indeed the only source of linguistic expressibility we have to take into consideration. Still, on the basis of the debate between Quine and Boolos on second-order logic, it might be tempting to argue that, if both set-theoretic and plural semantics model the language of second-order logic while providing very different models for it, then the modelling of (fragments of) natural language into a formal language is not *per se* indicative of what is expressed and ontologically presupposed by natural language.<sup>11</sup> This latter view is compatible with the claim that, as for the issue of the ontological commitment of quantification in formal languages, natural language can provide models that, at the very least, question views inspired by Quinean assumptions such as (1) and (2) from above.

**3. Reference** The issue I will be interested in, in this paper, concerns yet another fundamental notion, i.e. the notion of *reference*, in particular as for *singular terms* such as proper names or individual constants. My question concerns the nature of this notion, its functioning, and how it comes about, i.e. how it is fixed. My main claim will be that looking into natural language for answers to these questions is not particularly illuminating, since, I claim, the philosophical analysis of reference in natural language is usually carried out by notions that do not come exclusively from semantic considerations and, as such, contribute to blurring it. My main thrust will be that we can go look into formal languages, instead, and devise a form of reference,

---

<sup>9</sup> Quine (1956, p. 180).

<sup>10</sup> Especially examples of plural reference in natural language that is not formally reducible to first-order logic, but has to be translated into second-order logic, e.g. the so-called *Geach-Kaplan sentence* “Some critics admire only one another”.

<sup>11</sup> Boolos (1984, p. 449): “The lesson to be drawn from the foregoing reflections on plurals and second-order logic is that neither the use of plurals nor the employment of second-order logic commits us to the existence of extra items beyond those to which we are already committed. We need not construe second-order quantifiers as ranging over anything other than the objects over which our first-order quantifiers range, and, in absence of other reasons for thinking so, we need not think that there are collections of (say) Cheerios, in addition to the Cheerios. Ontological commitment is carried out by our *first-order* quantifiers; a second-order quantifier needn’t be taken to be a kind of first-order quantifier in disguise, having items of a special kind, collections, in its range. It is not as though there were two sorts of things in the world, individuals, and collections of them, which our first- and second-order variables, respectively, range over and which our singular and plural forms, respectively denote. There are, rather, two (at least) different ways of referring to the same things, among which there may well be many, many collections”.

which, unlike reference in natural language, is stripped to its semantic essentials, which is everything we need for reference to work in the first place. So, with respect to the proposed methodological divide, I am definitely on the side of formal languages in this. As Shapiro (1997) stresses,

Probably the most baffling, and intriguing, semantic notion is that of *reference*. (...). How does a term come to denote a particular object? What is the nature of the relationship between a singular term ('Fido') and the object that it denotes (Fido), if it denotes anything? (p. 139).

Again, we have at least two options here: either we go look into the semantics of formal languages or we go look into the semantics of natural language. At first glance, the first route might seem to lead us nowhere:

Notice that model theory, by itself, has virtually nothing to say on this issue. In textbook developments of model theory, reference is taken as an unexplicated *primitive*. It is simply *stipulated* that an 'interpretation' includes a function from the individual constants to the domain of discourse. This is a mere shell of the reference relation. (...) As far as the model-theoretic scheme goes, it does not matter how this reference is to be accomplished or whether it can be accomplished in accordance with some theory or other. (...) As far as model theory goes, reference can be *any* function between the singular terms of language and the ontology (*Ibid.*)

To sum up the quotation from Shapiro, it seems that looking into model-theory for illumination on what reference of singular terms is and how it works is hopeless: any function will do. So, we might as well turn to natural language, in order to see if we can do any better:

If we assume that ordinary languages are understood and if we accept the premise that model theory captures the structure of ordinary interpreted languages, then we can do better. There is, of course, no consensus on how reference to ordinary physical objects is accomplished. The theories are legion. [But at the very least we can presume that] reference to proverbial medium-sized physical objects is accomplished (...) (*Ibid.*, text in square brackets added.)

Nevertheless, even though, on the one hand, the semantics for formal languages might be thought to be of no help in analysing what the notion of reference is and how reference is fixed, on the other, at least two, to some extent incompatible views of reference have been proposed in the analysis of natural language, i.e. Kripke's direct reference view and the classical descriptivist theory of reference supported by Russell and, to some extent, Frege, so that it might be questioned whether the reference relation is captured by any of those views at all.

To this latter claim, it might be objected that, even if natural language is of no more help than the semantics of formal languages with respect to the analysis of reference, at the very least in natural language we have a sense of what seems to be going on.<sup>12</sup> Nevertheless, the attempted

---

<sup>12</sup> Notice that reference is taken as a primitive both in the semantics of formal languages and in (most of?) those philosophical pictures in philosophy of language. Reference is not reduced to more basic notions, but scholars have just tried to explain or elucidate it.

analysis of reference and how it is fixed in natural language usually brings influences from different areas of philosophy, not semantics alone, but also e.g. metaphysics or philosophy of mind. In particular, the notions of *individuation* from metaphysics and of *intentionality* from the debate in the philosophy of mind carry a rather crucial import in the analysis of reference in natural language. In this paper, I would like to focus on the impact of the notion of individuation on reference fixing, in particular as for the two classic views on reference in philosophy of language.

A word of caution, though. The notion of individuation may not be particularly transparent conceptually: Lowe (2003) distinguishes between a *metaphysical* interpretation of individuation, and an *epistemic* interpretation of it. By *metaphysical individuation*, Lowe means:

an ontological relationship between entities: what “individuates” an object, in this sense, is whatever it is that makes it the single object that it is – whatever it is that makes it *one* object, distinct from others, and the very object that it is as opposed to any other thing (Lowe, 2003, p. 75).

As Lowe argues, this conception of individuation may be very hard to capture. Epistemic individuation, i.e. the “singling out” of an entity as a “distinct object of perception, thought, or linguistic reference” (Lowe, 2003, p. 75), on the other hand, seems more straightforward, especially since it is often accompanied by some identity criterion.

In this respect, it seems reasonable to point out that the two main opposing views of reference in the philosophy of language seem to require a form of individuation for reference of singular terms to be fixed. Kripke’s initial baptism, in fact, can take place by pointing at an object and tagging it with a proper name – at least when the baptism takes place *in praesentia*, otherwise definite descriptions will do.<sup>13</sup> Also, according to the descriptivist theory of reference,<sup>14</sup> the fixing of reference takes place via definite descriptions, which are a means to single out objects of reference. So, either way, reference via proper names seems to presuppose epistemic individuation in that it requires that objects of reference are somehow singled out, in order for reference of proper names to be fixed. This feature of reference via proper names, shared by the views mentioned above, induces what I like to call *canonical reference*.<sup>15</sup>

#### 4. From Natural Language to Formal Languages

The idea that the fixing of reference of singular terms in natural language is usually accompanied by epistemic individuation is not problematic *per se*. Nevertheless, as soon as we ask for a uniform view of what reference is and how it works, and we take natural language as the only possible starting point in this enquiry, we might also be tempted to claim that the notion of reference in natural language is *the one and only* notion of reference. If, furthermore, we also want to apply this view of reference in natural language to the functioning of reference within formal frameworks, I think it is, first of all, disputable the notion of reference in natural language is to be reliably and exhaustively modelled in formal languages; secondly, implausible that, by extension, it is this kind of reference that takes place in formal languages. Hardly a function of assignment individuates, hardly by description (unless the function is provided with specific instructions) – let alone by ostension. If so, we are forced to assume the existence of at least two different notions of reference, i.e. one in natural language and the other in formal languages. But if we do think that the basic notions analytic philosophy

<sup>13</sup> See Kripke 1980, p. 96 & fn. 42.

<sup>14</sup> E.g. Russell’s and, to some extent, Frege’s.

<sup>15</sup> See also Boccuni & Woods (2018) for similar considerations.

tries to unveil would better be unique, this option seems useless. The alternative view, which is the one I support, is that there is only one basic notion of reference, such that non-semantic considerations are not necessary to fruitfully carry out its analysis. In particular, I do think that we can do only with one, basic notion of reference, which nevertheless does not presuppose individuation, and such that it can be fruitfully analysed in some appropriate semantics for formal languages. In a nutshell, my view is that, though reference is indeed a primitive tool of languages, it can still be stripped down of features that are, in some sense, a philosophical overload, leading thus to a deflated notion of reference i.e. *arbitrary reference*.<sup>16</sup> The outcome is an analysis of reference which is carried out in formal semantics, and such that it can also be reasonably applied to natural language.<sup>17</sup>

In some recent literature,<sup>18</sup> a case has been made for (at least some) singular terms (in mathematical discourse) to be interpreted as *parameters* or *arbitrary names*: terms grammatically behaving like singular terms, but, at least to some extent, semantically functioning as pronouns or indefinites. For instance, in mathematical discourse arbitrary names can be introduced by locutions of the form ‘Let  $n$  be an arbitrary natural number’. This move has the advantage of being proposed on semantic-linguistic (non-metaphysical) grounds, but can be specified in very different, even incompatible, ways. Brandom (1996) provides a survey of the possible options on the table, according to which we may develop two distinct accounts of reference in mathematical discourse.

According to the first account, one may take a liberal view and make a case for two primitive and independent notions of reference: the one I called *canonical reference*; and *arbitrary reference* as embodying some notion of reference other than canonical. This view has the obvious disadvantage of multiplying the notions of reference *præter necessitatem*. Furthermore, this view does not seem appealing, for at least a very general reason: it is hard to justify that we have two different notions of reference, if we have a hard time in capturing what reference should be in the first place. According to the second option presented by Brandom (1996), one may have a confrontational attitude on the issue of reference. This latter view, in turn, may be spelled out in two possible ways:

- (i) (at least some) singular terms are parameters, and they do not refer;
- (ii) singular terms genuinely refer, just not (necessarily) via individuation: what is really basic is arbitrary reference, and, if at all, canonical reference may be built up from it.

---

16 For a claim in a similar spirit, though arising from a possible semantics of proper names in natural language, see e.g. Cumming (2008).

17 It might be objected that my view on reference does not imply that the best methodology to tackle this issue requires to analyse it in a formal framework. After all, it would still be compatible with my view that arbitrary reference is a primitive tool of natural language, in which it is already understood, and formal frameworks at best model it, to the effect that the proposed methodological priority of formal over natural language may well be reversed. Nonetheless, unlike in formal frameworks, I find it hard to envisage a clear methodology within natural language to account for arbitrary reference. The only examples of primitive arbitrary reference in natural languages, in fact, seem to be indefinites such as ‘a natural number’. Not all linguists and semanticists agree on the correct interpretation of indefinites, though: indefinites might be treated either as quantified expressions (sometimes existentially, sometimes universally), or as primitively referential, or as essentially predicates with free variables – see, e.g., King & Lewis (2016) and Ludlow (2018) for surveys on these different readings. It seems, then, that the systematic ambiguity of natural language that caused Frege’s scorn is still alive and kicking some a hundred and forty years after the *Begriffsschrift* and regardless the enormous progress in philosophy and linguistics. As things stand, thus, I think it is still methodologically preferable to look into formal frameworks rather than natural language to account for arbitrary reference. I would like to sincerely thank an anonymous reviewer for pointing this out to me.

18 See e.g. Brandom (1996), Pettigrew (2008), Shapiro (2008, 2012), Woods (2014).



Pettigrew (2008) and Shapiro (2008, 2012) seem to pursue option (i), since they take arbitrary names to be non-referential. Applying this view seems uneffective, but I'll get back to it in §5. In what follows, I will pursue option (ii), namely the option by which we still have genuine referentiality, *contra* Shapiro and Pettigrew, and avoid appealing to individuation, thus supporting a metaphysically deflated notion of reference, i.e. *arbitrary reference*.<sup>19</sup>

**5. Arbitrariness and Genuine Reference, i.e. of Love and Logic**

The way in which I like to understand arbitrariness in arbitrary reference might be called the *epistemicist view*. Roughly, this view takes arbitrariness to be epistemic arbitrariness. We cannot *know which* individual an arbitrary term refers to, but the arbitrary term genuinely refers: it refers to a full-fledged individual, equipped with properties. The motivations for this view can be found in natural deduction. Breckenridge & Magidor (2012), Martino (2001, 2004), and Boccuni (2013) conceive of arbitrary reference as genuine reference, and the arbitrariness involved in it as an *epistemic* feature.<sup>20</sup> Mathematicians (and logicians, for that matter) often use locutions like 'Let  $n$  be an arbitrary natural number' in logico-mathematical discourse. These kinds of expressions introduce reference to, e.g., an arbitrary natural number  $n$ , and, according to Breckenridge & Magidor (2012), are governed by the following principle of arbitrary reference:

*Arbitrary Reference:* It is possible to fix the reference of an expression arbitrarily. When we do so, the expression receives its ordinary kind of semantic value, though we do not know and cannot know which value in particular it receives.

According to Martino (2001), the possibility of referring, at least in a non-canonical way, to any entity of a universe of discourse is presupposed both by logical and mathematical reasoning, even when non-denumerable domains are concerned. As a consequence, quantification itself logically presupposes this possibility of referring to each and every element of a domain, before we consider those elements through generalisation. Such a possibility of reference is very well expressed by the crucial role arbitrary reference plays both in formal and informal reasoning. Its cruciality lies in that arbitrary reference exhibits two different logical features that make it essential for performing proofs, i.e. *arbitrariness* and *determinacy*.

Through arbitrary reference, we may consider *any* entity  $a$  of a universe of discourse: as arbitrary reference does not single out specific entities, ' $a$ ' does not refer to an individuated entity  $a$ , rather it refers to an arbitrary  $a$ . Consequently, the epistemic ignorance about which individual  $a$  is, and thus of what its specific properties are, retains the general validity of the arguments about  $a$ . This feature provides a justification for the correct uses of the rule of universal introduction in natural deduction, which sanction as legitimate a deduction from  $\phi a$  to  $\forall x\phi x$  just in case the universal conclusion  $\forall x\phi x$  does *not* depend upon any assumption concerning  $a$ . This restriction is based on the rationale that, in order to legitimately conclude that all  $x$  are  $\phi$  from the fact that an arbitrary  $a$  is  $\phi$ , we have to assure that no specific properties of  $a$  intervene in the deductive step to  $\forall x\phi x$ . Those specific properties of  $a$ , in fact, may not be shared by all the individuals in the domain, and thus the soundness of the deductive step to  $\forall x\phi x$  might be at risk.

At the same time, though, an arbitrary name ' $a$ ' is used to refer to the *same* individual  $a$

---

<sup>19</sup> See Boccuni & Woods (2018) for the application of this view to the foundational programme in the philosophy of mathematics known as *Neologicism*.

<sup>20</sup> Of course, there is also Fine's view on arbitrary objects, but taking it into consideration would lead us too far astray, so I shall leave the comparison between my proposal and Fine's to some further work.

within a derivation on  $a$ . Again, a crucial reason for this is to be found in the requirement of soundness we want to impose on some valid argument schemas. If sameness of reference were not a basic ingredient of derivations, soundness would be in jeopardy. Consider the rule of existential elimination in natural deduction. When we pass from a premise of the form  $\exists x\phi x$  to the auxiliary assumption  $\phi(a)$ , ' $a$ ' has to be an unused arbitrary name, or at least it has not to appear in any of the assumptions which  $\exists x\phi x$  depends upon. Consider now the following (invalid) deduction:

- |     |                     |                    |
|-----|---------------------|--------------------|
| (1) | $\exists xHx$       | $\mathcal{A}$      |
| (2) | $\exists x\neg Hx$  | $\mathcal{A}$      |
| (3) | $Ha$                | $\mathcal{A}$      |
| (4) | $\neg Ha$           | $\mathcal{A}$      |
| (5) | $Ha \wedge \neg Ha$ | 3,4 intr. $\wedge$ |

Invalidity stems out from that, in eliminating the existential quantifiers respectively from (1) and (2), we use the very same arbitrary name in (3) and (4). Say that  $H$  is the property of being even and  $x$  varies over the natural numbers: (1) and (2) say, respectively, that there is at least a number which is even and there is at least a number which is not. Both these sentences are true in the standard model of Peano arithmetic. Nevertheless, if we use the same arbitrary name to perform existential elimination in the derivation above, in (3) and (4) we respectively say that a number is even and that *the very same number* is not, from which the contradiction in (5). For this reason, using an already used arbitrary name in (4) cannot be allowed.

In order to explain the invalidity of the derivation (1)-(5) ' $a$ ' must be referring to the *same*, though arbitrary, individual both in lines (3) and (4). Thus, in order to achieve soundness in the previous example, in line (4) we have to use a different arbitrary name than ' $a$ ', because we need to express that a *different* individual than  $a$  is  $\neg H$  within the same derivation, in accordance with the restrictions imposed by the rule of existential elimination. But then again, in order to distinguish between  $a$  and any other arbitrary individual that is  $\neg H$ , we have to assume that  $a$  is a *determinate*, though arbitrary, individual of the domain. The motivation for this requirement is very nicely explained by Suppes:

(...) ambiguous names,<sup>21</sup> like all names, cannot be used indiscriminately. The person who calls a loved one by the name of a *former* loved one is quickly made aware of this. (...) Such a happy-go-lucky naming process is bound to lead to error, just as we could infer a false conclusion from true facts about two individuals named 'Fred Smith' if we did not somehow devise a notational device for distinguishing which Fred Smith was being referred to in any given statement. The restriction which we impose to stop such invalid arguments is to require that when we introduce by existential specification an ambiguous name in a derivation, that name has not previously been used in the derivation (Suppes, 1999, p. 82).

The reasons for restricting the rules of introduction and elimination of quantifiers in natural deduction are semantic: in derivations, we perform a semantic reasoning that we want to proof-theoretically capture by deductive rules and the restrictions on them, in order to retain the soundness of our reasoning. Such a reasoning is crucially based, on the one hand, on arbitrariness and, on the other, on sameness and determinacy of arbitrary reference. But

---

21 I.e. arbitrary names.

then again, in order to make sense of sameness and determinacy, and consequently of the requirements we impose on deductive rules for the sake of soundness, we have to assume the genuine referentiality of arbitrary names to begin with. Genuine referentiality is a necessary condition for soundness.<sup>22</sup>

**6. First-Order Logic and a Semantics for Arbitrariness**

Let us now come to a possible semantics for arbitrary reference, and let us consider a formal framework that is as simple as possible: classical *first-order logic* (FOL). There are several semantics that may be utilised to make sense of arbitrary reference in this epistemic nuance, e.g., at least, choice functions, Hilbert-style  $\epsilon$ -terms, and more recently Martino's acts of choice semantics. Regardless what the preferred semantics is, any of them may be used to provide a semantics that fixes the reference of arbitrary names in classical FOL. I will provide a semantics in terms of choice functions, but I think nothing really substantial bears upon it. For the sake of simplicity, let us focus on the monadic fragment  $\mathcal{L}$  of the language of classical FOL, i.e. the fragment containing an infinite list of free variables  $\{x, y, z, \dots\}$ , an infinite list of arbitrary names  $\{a, b, c, \dots\}$ , and an infinite list of monadic predicate letters  $\{F, G, H, \dots\}$ . Let us also assume that the arbitrary names are the only singular terms, i.e. in  $\mathcal{L}$  there are no individual constants.

Let now  $D$  be a non-empty domain, and let  $f$  be a choice function of assignment taking the set of the singular terms of  $\mathcal{L}$  as its domain.<sup>23</sup> In order to model epistemic arbitrary reference, we want  $f$  to assign 'a' to a member of a set of candidate referents from  $D$ . Informally speaking, in order to attach a referent to 'a' we go look into the domain for possible referents, but since 'a' is arbitrary the set of such possible referents may not contain exactly one member: there might be different individuals in  $D$  that are eligible as denotations of an arbitrary name 'a';  $f$  picks exactly one of them, still we don't know which one. So, considering a subset  $C$  of  $D$  containing candidate referents for 'a',  $f$  assigns a member of  $C$  to 'a':

$$f: \{a \in \mathcal{L} : a \text{ is an arbitrary name}\} \rightarrow C \subseteq D,$$

such that:

- (1)  $C \subseteq D := \{i : i \in D \wedge \mathcal{M} \models \Phi\}$ , where  $i$  is an individual in  $D$  and  $\Phi$  is an arbitrary condition expressible in  $\mathcal{L}$ ;
- (2.a) if  $C \neq \emptyset$ , then  $f(a) = i \in C$ , for any  $i \in C$ ;
- (2.b) if  $C = \emptyset$ , then  $f(a) = f(D)$ .<sup>24</sup>

We have to extend this to predicate letters: assuming that predicate letters are interpreted

---

<sup>22</sup> Those who support the non-referentiality of arbitrary reference (e.g. Pettigrew, Shapiro) should provide some argument for explaining how classical formal and informal reasoning functions in the way it does, e.g., by certain constraints on introduction and elimination of quantifiers.

<sup>23</sup> Free variables are interpreted as usual, through arbitrary infinite sequences  $S_i$  of individuals of  $D$  that  $S_i$  correlates to free variables, for  $1 \leq i \leq \omega$ .

<sup>24</sup> One might wonder whether, in order to select appropriate candidate denotations for 'a', some form of individuation is still required, since we have to distinguish candidate referents for 'a', which are  $C$ , from those individuals in  $D$  that are  $\neg C$  and, thus, cannot be candidate referents for 'a'. Nevertheless, since the underlying first-order logic is classical and also complete, any subset  $C$  of  $D$  will be such that its complement  $\neg C$  is completely determined, regardless our epistemic ability to single out which objects, if any, are in  $C$  and which are not. Since the logic is classical, we know for a fact that, if any individual at all is in  $C$ , it will not be in  $\neg C$ , and that both these subsets of  $D$  are completely determined with respect to the individuals they contain, even though we might not be able to single those individuals out. I would like to thank an anonymous reviewer for pointing this out to me.

objectually, they are referential as well. Let  $g$  be a choice function of assignment taking the set of all predicate letters of  $\mathcal{L}$  as its domain, and ' $F$ ' a predicate letter. The function  $g$  assigns ' $F$ ' to a member of a class of subsets of  $D$  that are candidate referents for ' $F$ '. Informally speaking, in order to attach a referent to ' $F$ ' we go look into a class of subsets of  $D$  for possible referents, but since ' $F$ ' is interpreted arbitrarily such a class may not contain exactly one member: there might be different subsets of  $D$  that are eligible as denotations of ' $F$ ';  $g$  picks exactly one of them, still we don't know which one. So, being  $\pi(D)$  a class of subsets of  $D$  containing candidate referents for ' $F$ ',  $g$  assigns a member of such class to ' $F$ ':

$$g : \{F \in \mathcal{L} : F \text{ is predicate letter}\} \rightarrow \pi(D),$$

such that

- (3)  $\pi(D) := \{X : X \subseteq D \wedge \mathcal{M} \models \psi X\}$ , where  $X$  is a member of  $\pi(D)$  and  $\psi$  is an arbitrary condition expressible in  $\mathcal{L}$ ;
- (4.a) if  $\pi(D) \neq \emptyset$ , then  $g(F) = C \in \pi(D)$ , for any  $C \in \pi(D)$ ;
- (4.b) if  $\pi(D) = \emptyset$ , then  $g(F) = g(\pi(D))$ .

It might be objected that predicate letters in FOL are constants, functioning as proper names of properties of individuals, and thus my proposed view is not faithful to their intended meaning. Replying this objection is connected to one of the issues mentioned in passing in §4, bullet point (ii), i.e. the possible construal of canonical reference from arbitrary reference. If canonical reference is to be recovered, we can do so by taking arbitrary reference as primitive, and then build canonical reference from it. What is needed is some formal device for expressing individuating conditions, likely in their epistemic nuance i.e. by conditions singling out a specific object of reference. A possible way to do this would be to provide a uniqueness condition by a definite description. The uniqueness condition would impose a constraint on the codomain of a choice function of assignment, i.e. it would work in the limit case where the codomain of the function contains exactly one member. This goes as for both predicate letters such as ' $F$ ' and the introduction of individual constants from arbitrary names.

We can also extend this strategy to the semantics of natural language. All in all, when a Kripkean initial baptism takes place, reference fixing is usually accompanied by some individuating conditions, and this can be modelled as I just suggested. Though, that *de facto* reference fixing in natural language goes with individuation does not imply that the fixing of reference of a proper name and its semantics are necessarily and *a priori* intertwined with metaphysics.

A somewhat more difficult question concerns whether a *general* theory of arbitrary reference can be provided, such that it is applicable to the semantics of any formal theory, whether logical or not. After all, I only provided a semantics for classical first-order logic, and this might raise qualms about the feasibility of extending such arbitrary reference semantics to more complex formal languages. In principle, I see no obstacle to this, since reference is indeed a basic semantic tool of any language, but this topic I shall leave for future research.

In this paper, I investigated one of the most basic notions in the analytic debate, i.e. reference (of singular terms). My aim was to tackle the notion of reference in the light of a classic methodological divide: whether its analysis, like the analysis of other basic notions, should be carried out in natural language or in the semantics of formal frameworks. I argued that the most sensible strategy is to rely on formal resources, because relying on natural language

## 7. Concluding Remarks

implies a philosophically overloaded analysis, weighed down with notions from philosophical areas other than semantics, like e.g. metaphysics. I argued for a deflated notion of reference, namely arbitrary reference, as a genuine, primitive notion, and provided a semantics for it apt for classical first-order logic – the simplest formal framework in which the referentiality of singular terms is relevant.

#### REFERENCES

- Boccuni, F. (2013). Plural Logicism. *Erkenntnis*, 78(5), 1051-1067;
- Boccuni, F., & Woods, J. (2018). 'Structuralist Neologicism'. *Philosophia Mathematica*, nky017, <https://doi.org/10.1093/phimat/nky017>;
- Boolos, G. (1984). To be is to be the Value of a Variable (or the Values of Some Variables). *Journal of Philosophy*, 81, 430-450;
- Boolos, G. (1985). Nominalist Platonism. *Philosophical Review*, 94, 327-344;
- Brandom, R. (1996). The Significance of Complex Numbers for Frege's Philosophy of Mathematics. *Proceedings of the Aristotelian Society*, 96, 293-315;
- Breckenridge, W., & Magidor, O. (2012). Arbitrary reference. *Philosophical Studies*, 158(3), 377-400;
- Cook, R. (2013). *Paradoxes*. Oxford: Wiley Blackwell;
- Cumming, S. (2008). Variabilism. *Philosophical Review*, 117(4), 525-554;
- King, J., & Lewis, K. (2016). Anaphora. *Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/entries/anaphora/ProbAnap>;
- Kripke, S. (1975). Outline of a theory of truth. *Journal of Philosophy*, 72(19), 690-716;
- Lowe, E. J. (2003). Individuation'. In M. J. Loux & D. W. Zimmerman (Eds.), *The Oxford Handbook of Metaphysics* (pp. 75-95). Oxford: Oxford University Press;
- Ludlow, P. (2018). Descriptions. *Stanford Encyclopedia of Philosophy*. Retrieved from <https://plato.stanford.edu/entries/descriptions/IndDes>;
- Martino, E. (2001). Arbitrary Reference in Mathematical Reasoning. *Topoi*, 20, 65-77;
- Martino, E. (2004). Lupi, pecore e logica. In M. Carrara & P. Giaretta (Eds.), *Filosofia e logica* (pp. 103-133). Catanzaro: Rubbettino;
- Pettigrew, R. (2008). Platonism and Aristotelianism in Mathematics. *Philosophia Mathematica*, 16(3), 310-332;
- Priest, G. (2006). *In Contradiction*. Oxford: Oxford University Press;
- Quine, W.V.O. (1956). Quantifiers and Propositional Attitudes. *Journal of Philosophy*, 53, 177-187;
- Quine, W.V.O. (1986). *Philosophy of Logic*. Cambridge: Harvard University Press. Revised Edition;
- Shapiro, S. (1997). *Philosophy of Mathematics*. New York, Oxford: Oxford University Press;
- Shapiro, S. (2008). Identity, Indiscernibility, and Ante Rem Structuralism: The Tale of *i* and *-i*. *Philosophia Mathematica*, 16(3), 285-309;
- Shapiro, S. (2012). An "i" for an *i*: Singular Terms, Uniqueness, and Reference. *Review of Symbolic Logic*, 5, 380-415;
- Woods, J. (2014). Logical Indefinites. *Logique et Analyse*, 227, 277-307.

---

BIANCA CEPOLLARO

Università Vita-Salute San Raffaele  
cepollaro.bianca@univr.it

---

# INTUZIONI LINGUISTICHE E FILOSOFIA SPERIMENTALE: METODI A CONFRONTO

---

## abstract

*Nell'ultimo decennio si è infittito il dibattito sul ruolo della filosofia sperimentale rispetto alla cosiddetta filosofia 'in poltrona' ('armchair philosophy'). La discussione, che riguarda i metodi della filosofia analitica in generale, ha ricadute di grande interesse per la filosofia del linguaggio in particolare. In questo articolo presento un caso in cui una questione centrale nello studio dei termini espressivi – cioè se il contenuto offensivo delle espressioni denigratorie sopravviva o no nell'interazione con operatori semantici e in particolare con la negazione – è stata affrontata quasi contemporaneamente adottando metodi filosofici differenti: (i) in poltrona, (ii) attraverso l'applicazione di test linguistici e (iii) attraverso metodi sperimentali. I tre metodi – così diversi per tradizione e prospettive – offrono risultati convergenti. Casi come questi mostrano come una pluralità metodologica e un atteggiamento più ecumenico rispetto allo scontro tra filosofia sperimentale e filosofia 'in poltrona' non siano da temere ma anzi da perseguire.*

---

## keywords

*metodi, filosofia sperimentale, armchair philosophy, intuizioni linguistiche*

**1. Introduzione** Nell'ultimo decennio, si è infittito il dibattito intorno al ruolo delle *intuizioni*, da intendersi sia come le intuizioni dei filosofi di professione, sia come le intuizioni dei cosiddetti 'lays', cioè la gente comune, o non-filosofi. Le intuizioni che ricoprono maggior interesse in questa discussione sono quelle relative agli esperimenti mentali (si vedano in proposito Weinberg *et al.*, 2001; Williamson, 2007; Swain *et al.*, 2008; Liao, 2008; Machery, 2017).

Gli interrogativi che hanno animato questo dibattito riguardano l'affidabilità delle intuizioni in generale, l'idea che le intuizioni degli esperti siano più affidabili di quelle dei non-esperti, la difficoltà nello stabilire chi siano gli esperti in ciascun campo, e così via. Inoltre, il fatto che su innumerevoli questioni filosofiche permanga un profondo disaccordo tra i filosofi di professione potrebbe suggerire che le intuizioni sugli esperimenti mentali di questi ultimi non sono in fondo così affidabili: se così fosse, dovremmo osservare nel tempo una certa convergenza di opinioni riguardo agli esperimenti mentali considerati dalla comunità. Ad alimentare la sfiducia nell'attendibilità delle intuizioni, vi sono anche studi che mostrano come fattori quali istruzione, situazione socioeconomica, formulazione del problema, forse genere e così via siano in grado di influenzare le intuizioni sugli esperimenti mentali (Weinberg *et al.*, 2001).

La discussione sull'affidabilità teorica delle intuizioni sugli esperimenti mentali, insieme alla questione su chi conti come 'esperto' in materia di problemi filosofici, ha conseguenze rilevanti sul ruolo della filosofia sperimentale rispetto alla più tradizione filosofia 'in poltrona'. Vi sono posizioni radicali a favore dell'uno o dell'altro *modo* di fare filosofia, così come posizioni più moderate (si vedano per esempio Kauppinen, 2007 e Knobe & Nichols, 2013), ma l'interrogativo da cui prendo le mosse è se questo dibattito possa davvero interessare solo chi è impegnato nell'indagine sui metodi della filosofia analitica, o se non sia piuttosto una questione con cui filosofe e filosofi si debbano confrontare nel portare avanti le proprie ricerche. Ogni modo di fare filosofia assume, quand'anche implicitamente, una certa posizione metafilosofica: questa osservazione scontata basta di per sé a rendere la domanda sui metodi una domanda di tutti. Il caso che in questo articolo discuto riguarda una questione molto circoscritta all'interno del dibattito sugli espressivi in filosofia del linguaggio e in linguistica. Quello che vorrei fare dopo aver fornito delle coordinate di massima sulla materia di discussione (Sezione 2) è mostrare quali metodi differenti siano stati adoperati per dipanare il dilemma in esame: dalla cosiddetta *armchair philosophy*, cioè la filosofia tradizionale che affronta i problemi filosofici sviluppando nuove analisi senza bisogno di collezionare dati empirici (Sezione 3.1), all'impiego di test linguistici per scandagliare le intuizioni dei parlanti – non necessariamente dei filosofi – (sezione 3.2), fino ad arrivare ai metodi di filosofia

sperimentale (Sezione 3.3). Il mio scopo è presentare un caso virtuoso in cui i diversi approcci messi in campo in filosofia hanno fornito risultati congruenti.

Gli *slurs*, in italiano ‘epiteti denigratori’, sono termini che denigrano i propri bersagli in quanto appartenenti a una certa categoria, spesso chiamata ‘categoria target’. Tali categorie possono essere l’orientamento sessuale, l’etnia, la provenienza geografica, la religione, eccetera. Alcuni esempi di *slurs* in italiano sono espressioni come ‘frocio’, ‘negro’, ‘terrone’, e così via. Vari nuclei di interesse hanno portato questi termini al centro del dibattito filosofico negli ultimi anni: innanzitutto essi sembrano mettere in discussione la distinzione fatti/valori, dal momento che sono termini che contemporaneamente veicolano un’informazione fattuale ed esprimono un giudizio valutativo (negativo); oltre a ciò, lo studio degli epiteti denigratori diventa particolarmente rilevante quando ci si interroga su questioni legate al linguaggio d’odio e alla sua permissibilità nei discorsi pubblici, privati e sui social; inoltre, gli *slurs* sono spesso analizzati non solo come termini che rispecchiano un certo atteggiamento discriminatorio, ma come strumenti di propaganda (Cepollaro, 2017, p. 62 segg) e secondo alcuni – si veda ad esempio Jeshion, 2013a; 2013b – gli *slurs* svolgono una vera e propria funzione di disumanizzazione dei loro bersagli; infine – ed è su questo che ci soffermeremo in questo lavoro – gli epiteti denigratori presentano un comportamento linguistico particolare. La domanda a cui molti in filosofia e linguistica hanno cercato di rispondere riguarda la *relazione* tra i termini denigratori che chiamiamo ‘*slurs*’ e quei contenuti dispregiativi a cui essi sono tipicamente associati. In altre parole, la questione è se il contenuto denigratorio degli *slurs* sia parte del significato vero-condizionale degli epiteti stessi, se sia un’implicatura conversazionale, una presupposizione, un’implicatura convenzionale, un effetto perlocutorio, un effetto della forza illocutoria degli enunciati in cui gli *slurs* compaiono, e così via. Non esiste al momento una teoria accettata all’unanimità, ma possiamo se non altro affermare che la maggior parte degli autori e delle autrici sostiene che la componente denigratoria degli *slurs* *non* sia parte del loro significato vero-condizionale (si vedano per esempio Schlenker, 2007; Richard, 2008; Predelli, 2010; Camp, 2013; Jeshion, 2013a, 2013b; Bolinger, 2015; Cepollaro & Stojanovic, 2016). A motivare questa posizione, vi è l’osservazione di come la componente denigratoria di queste espressioni si comporta quando interagisce con gli operatori vero-condizionali. Si considerino i seguenti enunciati:

- (1) Lucia è terrona.
- (2) Lucia non è terrona.
- (3) Lucia è terrona?
- (4) Lucia potrebbe essere terrona.
- (5) Se Lucia è terrona, lo è anche suo fratello.
- (6) Non ci sono terroni.

Il contenuto dispregiativo veicolato da ‘terrona’ in (1) non sparisce quando il termine viene negato in (2) e (6), né quando il termine compare in una domanda (e dunque non è predicato del soggetto) in (3), né, ancora, quando si trova nell’ambito di un predicato modale in (4), o di un condizionale in (5). Come accennato poc’anzi, la maggior parte delle filosofe e dei filosofi coinvolti in questo dibattito ritiene che gli esempi appena proposti dovrebbero spingerci a scartare una teoria vero-condizionale della componente denigratoria degli epiteti: se questa componente fosse parte del significato vero-condizionale (spesso chiamato impropriamente ‘semantico’), allora essa dovrebbe interagire con gli operatori semantici in modo da, ad esempio, scomparire quando (1) viene negato o messa sotto forma di domanda, eccetera. Contro questa opinione diffusa, Robert May e Christopher Hom (si vedano Hom, 2008, 2010,

## 2. Il dibattito sugli *slurs*



2012; Hom & May, 2013, 2014, 2018) hanno sviluppato e difeso la teoria vero-condizionale. Parte degli argomenti dispiegati per sostenere tale teoria si appella a intuizioni che sembrano differire da quelle sostenute dalla maggior parte della comunità linguistica e filosofica. In particolare, secondo Hom e May, gli enunciati (2)-(6) non sono davvero ‘denigratori’ nello stesso senso in cui (1) può esserlo. Per i due studiosi, (2)-(6) non predicano proprietà dispregiative di nessuno e pertanto non sono denigratori in senso stretto. Anzi, (2) e (6) possono essere usati secondo Hom e May proprio per fare affermazioni *non* razziste. Di fronte a questo conflitto di intuizioni circa l’interpretazione di frasi come (2)-(6) – veicolano un contenuto dispregiativo nei confronti dei meridionali oppure no? –, l’obiettivo che ci si pone è di stabilire (i) quale interpretazione sia corretta (ammesso che via sia una interpretazione corretta) e (ii) cosa abbia suggerito l’interpretazione opposta. Per rispondere a questa domanda, ci troveremo a integrare strategie provenienti dalla cosiddetta filosofia ‘in poltrona’ (Sezione 3.1.), ma anche test linguistici (Sezione 3.2.) e infine dati dalla filosofia e linguistica sperimentali (Sezione 3.3).

### 3. Tre strategie

Vari autori e autrici (Anderson & Lepore, 2013, citati sotto, ma anche Jeshion, 2013a e Bolinger, 2015) hanno preso le distanze dalle intuizioni avallate da Hom e May circa la non-offensività degli enunciati negativi, classificando l’interpretazione non denigratoria di frasi tipo ‘Lucia non è terrone’ come metalinguistica.

#### 3.1. Dalla poltrona

Not all theorists agree that the negation of a slurring sentence is derogatory (e.g. Blackburn, p. 148; Dummett 2007, p. 527; Hornsby, p. 129, Hom, p. 31), invoking an alleged non-derogatory use of (7) in their favor.

(7) There are no niggers.

Careful examination of this datum is supposed to establish it as a case of metalinguistic negation (Horn, Ch. 6), as in (8).

(8) John is not good; he’s great.

(8), under a meta-linguistic reading, does not ascribe greatness while denying goodness. It registers that, for one reason or another, its speaker refuses to use ‘good’ for John. By analogy, should we conclude that what (9) (and by extension (7)) expresses is a meta-linguistic claim, namely, a refusal to apply the slur ‘nigger’ to anyone?

(9) There are no niggers; there are only African-Americans.

Blackburn says of such sentences they are a way of ‘disowning the attitude’. (Anderson & Lepore, 2013, p. 28)

Quella di Anderson & Lepore riportata sopra è, nelle loro stesse parole, un’*analogia* col fenomeno della negazione metalinguistica investigato da Horn in alcuni lavori fondamentali (Horn, 1985; 1989). Horn distingue due tipi di negazione che chiama ‘proposizionale’ e ‘metalinguistica’: la prima si applica al contenuto vero-condizionale del termine negato; la seconda serve a prendere le distanze da una certa scelta lessicale, da un tono, da una pronuncia.

A device for objecting to a previous utterance on any grounds whatsoever - including its conventional or conversational implicata, its morphology, its style or register, or its phonetic realization (Horn, 1985, p. 121).

It is relevant that metalinguistic negation can be employed by a speaker who wishes to reject the bigoted or chauvinistic point of view embodied in an earlier statement within the discourse context: (c) I beg your pardon: Lee isn’t an ‘uppity {nigger/broad/kike/wop/...}’ - (s)he’s a strong, vibrant {black/woman/Jew/Italian/...}. (...) ‘I’m not ‘colored’

- I'm black!', 'I'm not a 'gentleman of the Israelite persuasion' - I'm a Jew!' (Horn, 1985, p. 133).

Nella frase 'Lucia non è felice' la negazione non segnala la falsità di 'Lucia è felice', ma sottolinea che l'uso del predicato 'felice' non è adeguato a descrivere lo stato di Lucia (e questo sarebbe ancora più chiaro se per esempio proseguissimo così: 'Lucia non è felice, è felicissima'). Similmente, sostengono Anderson e Lepore, quando si leggono frasi come (2) o (6) come *non* offensive, non si sta interpretando la negazione come un operatore che nega un certo contenuto proposizionale, ma come un segnale che l'uso del predicato 'terrone' è inappropriato. Questa analogia, per quanto possa essere convincente, non è un argomento sufficientemente forte da scardinare la tesi di Horn & May. Nelle sezioni che seguono, consideriamo due modi diversi di testare questa ipotesi.

Una strategia per mostrare che effettivamente l'interpretazione non denigratoria di enunciati quali (2) e (6) dipenda da un'interpretazione metalinguistica della negazione è applicare appositi test linguistici. Questa procedura dovrebbe offrire dei risultati più solidi di una mera analogia 'dalla poltrona', nonostante anche in questi casi si faccia appello alle intuizioni linguistiche degli studiosi. Presento brevemente la strategia proposta in Cepollaro e Thommen, 2019, sezione 3.

### 3.2. Test linguistici

Nei suoi lavori fondamentali sulla negazione metalinguistica, Horn introduce dei test che consentono di distinguere la negazione metalinguistica da quella proposizionale. Uno di questi test prevede l'incorporazione morfologica della negazione (Horn 1985, 140). Si prendano i seguenti enunciati:

- (10) Lucia non è felice.
- (10') Lucia non è felice. [è triste]
- (10'') Lucia non è felice. [è felicissima]

Secondo Horn, (10) è ambiguo tra due interpretazioni che dipendono da due letture alternative della negazione. Se la negazione si legge come proposizionale, allora (10) sembra dire che Lucia è triste (come suggerito dalla parentetica in 10'). Se la negazione si legge come metalinguistica, allora (10) sembra dire che non è adeguato dire che Lucia è (solo) felice, perché è addirittura felicissima (10''). Horn suggerisce che per stabilire se la negazione in (10) sia metalinguistica o proposizionale, si può osservare cosa accade quando la negazione viene incorporata morfologicamente nel predicato 'felice', ottenendo il predicato 'infelice'. L'incorporazione infatti blocca le letture metalinguistiche della negazione e consente invece quelle proposizionali. Consideriamo i seguenti enunciati:

- (11) Lucia è infelice.
- (11') Lucia è infelice. [è triste]
- (11'') Lucia è infelice. [è felicissima]

Osserviamo con Horn che dicendo 'Lucia è infelice', abbiamo sì a disposizione l'interpretazione secondo cui Lucia è triste, ma non è più disponibile l'interpretazione secondo cui il termine 'felice' è inadeguato poiché Lucia è felicissima. In altre parole, l'incorporazione della negazione lascia sopravvivere solo l'interpretazione della frase coerente con l'uso proposizionale della negazione e fa scomparire le letture metalinguistiche. In Cepollaro (2017, pp. 102-114) e Cepollaro e Thommen (2019, sezione 3) il test dell'incorporazione della negazione viene applicato agli slurs: se la lettura non denigratoria di una frase come 'Lucia non è terrona' non

è più accessibile quando la negazione viene incorporata morfologicamente, questo vuol dire che – come sostengono Anderson & Lepore 2013; Jeshion 2013 e Bolinger 2015 – l’uso della negazione è metalinguistico. Viceversa, se la lettura non denigratoria di ‘Lucia non è terrona’ è ancora disponibile quando la negazione è incorporata morfologicamente, questo vuol dire che la negazione era proposizionale, come sostengono Hom e May. Questo interrogativo apparentemente molto circostanziato su come interpretare la negazione in (2) e (6) ha delle conseguenze teoriche più generali: una delle prove che Hom e May adducono per sostenere la teoria vero-condizionale degli epiteti denigratori è proprio che in frasi come (2) e (6) la negazione (proposizionale) può bloccare a volte il contenuto peggiorativo di questi termini. Tornando al test per la negazione metalinguistica, bisogna specificare che applicare questo test al caso degli *slurs* presenta una difficoltà: a differenza della coppia ‘felice’-‘infelice’, per incorporare morfologicamente la negazione in uno slur siamo costretti a creare un neologismo. Questo rende più complessa la valutazione delle intuizioni linguistiche e tuttavia non è impossibile ragionare su casi in cui si adoperano parole inventate: immaginiamo che un razzista che usi il termine ‘terrone’ per indicare i meridionali usi il termine ‘non-terrone’ per indicare tutti coloro che *non* sono meridionali. Consideriamo ora la frase ‘Lucia non è terrona’. Horn direbbe che (12) è ambigua tra una interpretazione della negazione come proposizionale (12’) e metalinguistica (12’’) (in entrambi i casi, la parentetica serve a suggerire una certa lettura della frase):

- (12) Lucia non è terrona.  
 (12’) Lucia non è terrona. [è di Milano]  
 (12’’) Lucia non è terrona. [è meridionale]

Per verificare se l’interpretazione non denigratoria di (12) dipenda da una lettura metalinguistica della negazione, incorporiamo la negazione nel predicato ‘terrone’, ottenendo ‘non-terrone’ e vediamo se, una volta incorporata la negazione e bloccata quindi ogni lettura metalinguistica, sia ancora possibile interpretare (12) come non peggiorativo. Se, quando la lettura metalinguistica della negazione è impossibile, il proferimento perde la sua lettura non denigratoria, allora possiamo inferire che la lettura non denigratoria altro non era che una lettura metalinguistica. Consideriamo ora (13) e le sue varianti.

- (13) Lucia è non-terrone.  
 (13’) Lucia è non-terrone. [è di Milano]  
 (13’’) Lucia è non-terrone. [è meridionale]

Sembrerebbe che (13’) sia ancora felice, cioè che sia possibile proferire (13) con l’intenzione comunicativa suggerita dalla parentesi in (13’); al contrario, (13’’) sembra infelice, come se non fosse possibile proferire (13) con l’intenzione comunicativa suggerita dalla parentesi. Se così fosse, allora la lettura non denigratoria di (12), che sparisce quando la negazione viene incorporata, dipende effettivamente da una lettura metalinguistica della negazione e tale lettura non denigratoria di (12) non conta come prova a favore dell’idea che sia possibile negare gli *slurs* eliminando la loro componente denigratoria (ipotesi invece accolta da Hom e May, interessati a sostenere che la componente dispregiativa degli *slurs* è parte del loro significato vero-condizionale). Per concludere, la strategia dei test linguistici, avanzata da Cepollaro e Thommen, 2019, sezione 3, suggerisce che, *contra* Hom e May, la negazione non interagisce col contenuto denigratorio degli *slurs*.

**3.3. Esperimenti** Panzeri & Carrus (2016) hanno perseguito un’altra strada per verificare quel che Anderson & Lepore (2013), Jeshion (2013) e Bolinger (2015) hanno suggerito ‘dalla poltrona’, conducendo

degli esperimenti (nella fattispecie, dei questionari). Il primo studio sperimentale riguarda, tra le altre cose, il modo in cui i parlanti valutano gli *slurs* in isolamento e gli *slurs* in frasi complesse. Nella prima parte del primo studio, Panzeri e Carrus chiedevano ai partecipanti di valutare su una scala crescente da 1 a 7 l'offensività di una serie di termini (tra cui alcuni *slurs*) in isolamento. Nella seconda parte del primo studio, chiedevano di valutare su una scala da 1 a 7 una serie di frasi complesse come frasi negative, frasi condizionali e domande, che in una condizione contenevano uno *slur* (es. 'Lino non è un frocio', 'Lino è un frocio?').

I risultati mostrano che il punteggio di offensività ottenuto dalle domande e dai condizionali contenenti *slurs* è in linea col punteggio ottenuto dagli *slurs* in isolamento. Non sorprende: come si accennava nella Sezione 2, la maggior parte di studiose e studiosi (con l'eccezione di Hom e May) ritiene che il contenuto denigratorio degli *slurs* resista all'incassamento (cioè all'interazione con gli operatori semantici). Quel che invece sorprende, e che sembra andare contro questa intuizione così diffusa, è che il punteggio di offensività ottenuto dalle frasi negative contenenti *slurs* è più basso di quello ottenuto dagli *slurs* in isolamento. Il risultato è interessante non solo perché contraddice quanto sostenuto nella letteratura filosofica e linguistica, ma anche perché sembra tracciare una distinzione tra la negazione e gli altri operatori semantici come ad esempio il condizionale. Questa apparente discrepanza – cioè l'idea che il contenuto denigratorio degli *slurs* interagisca con la negazione ma non, ad esempio, col condizionale – ha fatto nascere il sospetto che nel caso della negazione vi potesse essere qualche meccanismo particolare in azione. Nel secondo studio di Panzeri & Carrus (2016), gli autori hanno avanzato l'ipotesi secondo cui la ragione per cui le frasi negative contenenti *slurs* ottengono un punteggio più basso sia degli *slurs* in isolamento sia dei condizionali e delle interrogative contenenti *slurs* è che la negazione viene interpretata come metalinguistica. Per verificare tale ipotesi, gli autori sono ricorsi alla seguente tecnica: hanno fornito delle frasi negative (alcune delle quali contenevano *slurs*) e hanno chiesto ai parlanti di proporre una continuazione in modo che la frase risultante fosse dotata di senso. Per esempio, la frase 'Non piove a Milano' può essere continuata in modi diversi:

- (14) Non piove a Milano.
- (15) Non piove a Milano, c'è il sole.
- (16) Non piove a Milano, diluvia.

Il tipo di continuazione in (15) e (16) rivela modi diversi di interpretare la negazione in (14). La continuazione proposta in (15) rivela un'interpretazione proposizionale della negazione che agisce sul contenuto vero-condizionale di 'piove': quando c'è il sole, è falso che piova; in (16), invece, la continuazione rivela un'interpretazione metalinguistica della negazione: non si nega davvero che piova, ma si nega che 'piove' sia una descrizione adeguata, visto che addirittura diluvia. L'idea di Panzeri e Carrus è quindi di usare la tecnica della continuazione (cioè chiedere ai parlanti di proporre una continuazione in modo che la frase risultante sia dotata di senso) per verificare se la negazione in frasi come 'Lino non è un frocio' venisse interpretata proposizionalmente (come in (15)) o metalinguisticamente (come in (16)). Le risposte fornite dai soggetti sono state suddivise in risposte che suggeriscono un'interpretazione proposizionale della negazione (continuazioni tipo quelle in (17)-(18)) e risposte che lasciano intendere un'interpretazione metalinguistica della negazione (continuazioni come quelle in (19)-(20)):

- (17) Lino non è un frocio, ha una ragazza.
- (18) Lino non è un frocio, è solo effeminato.
- (19) Lino non è un frocio, è gay.
- (20) Lino non è un frocio, non mi piace che usi questo linguaggio omofobo.

Quel che Panzeri & Carrus (2016) hanno trovato è che la negazione nelle frasi contenenti slurs riceve una continuazione metalinguistica circa nel 33% dei casi raccolti. Questo dato può spiegare perché il punteggio di offensività attribuito alle frasi negative che contengono slurs sia inferiore a quello attribuito agli slurs in isolamento e agli altri tipi di frasi complesse contenenti slurs: i soggetti interpretavano la negazione in modo metalinguistico. Torniamo ora alla domanda da cui eravamo partiti, cioè se Hom e May abbiano ragione di sostenere che in frasi come ‘Lucia non è terrona’ la negazione blocchi il contenuto denigratorio degli slurs, rendendo gli enunciati non offensivi. Lo studio di Panzeri e Carrus sembra suggerire che l’intuizione ‘in poltrona’ di Anderson, Bolinger e Jeshion e supportata dai test linguistici applicati in Cepollaro (2017, pp. 102-114) e Cepollaro e Thommen (2019, sezione 3), sia corretta: il motivo per cui una frase in cui lo slur è negato, come ‘Lucia non è terrona’, può suonare non dispregiativa dipende da un’interpretazione metalinguistica della negazione. In questa sede, non entrerà nel merito delle altre strategie dispiegate da Hom e May per difendere ulteriormente la loro teoria vero-condizionale degli slurs, come ad esempio l’introduzione della nozione di ‘offesa’, che si aggiunge e si distingue da quella di ‘denigrazione’. Il centro di interesse qui non è stabilire se la teoria vero-condizionale degli epiteti sia in generale difendibile o no; la domanda a cui, con tre diverse metodologie, studiosi e studiosi hanno cercato di rispondere è molto più puntuale – e forse meno ambiziosa – e riguarda la possibilità di sostenere che vi siano delle letture non dispregiative di enunciati in cui gli slurs compaiono incassati (per esempio, gli enunciati considerati nella Sezione 2, (2)-(6)) perché gli operatori semantici interagiscono con il contenuto denigratorio degli slurs. Questa domanda, meno ambiziosa di quella sulla validità della teoria vero-condizionale, sembra ora avere una risposta negativa, dal momento che tre metodologie filosofiche indipendenti hanno fornito verdetti convergenti.

#### 4. Una conclusione ecumenica

In questo articolo si è considerato il dibattito sugli slurs e in particolare il modo in cui la negazione interagisce o non interagisce col contenuto denigratorio di queste espressioni. Questo caso costituisce un esempio di come diverse strategie metodologiche della filosofia analitica (dalle ipotesi dalla poltrona, alle applicazioni di test linguistici, alla ricerca sperimentale) possano integrarsi e offrire risultati coerenti e convergenti. Nell’Introduzione, ci si chiedeva se certe questioni di metodo (ad esempio l’uso che si fa delle intuizioni in filosofia e del valore privilegiato che si attribuisce alle intuizioni – anche linguistiche – dei filosofi) possano preoccupare solo chi si occupa di metafilosofia o se invece queste vadano affrontate da ciascuna/o filosofa/o nel proprio ambito di ricerca. Senza avere una risposta alla questione spinosa di come affrontare il problema del metodo in ogni ambito della filosofia analitica, un possibile approccio che questo caso concreto esemplifica è quello di integrare metodologie diverse, adottando un atteggiamento in un certo senso ecumenico rispetto allo scontro tra filosofia sperimentale e filosofia ‘in poltrona’, osservando se e quando queste diverse metodologie offrono risultati convergenti.

#### REFERENCES

- Anderson, L., & Lepore, E. (2013). Slurring words. *Nous*, 47 (1), 25-48;  
 Bolinger Jorgensen, R. (2015). The Pragmatics of Slurs. *Noûs*, 50 (3);  
 Camp, E. (2013). Slurring Perspectives. *Analytic Philosophy*, 54 (3), 330-349;  
 Cepollaro, B. (2017). *The Semantics and Pragmatics of Slurs and Thick Terms*. Tesi di dottorato, SNS Pisa, ENS Parigi. <https://tel.archives-ouvertes.fr/tel-01508856/> ;  
 Cepollaro, B., & Stojanovic, I. (2016). Hybrid evaluatives: in defense of a presuppositional account. *Grazer Philosophische Studien*, 93 (3), 458-488;

- Cepollaro, B., & Thommen, T. (2019). What's wrong with truth-conditional account of slurs. *Linguistics and Philosophy*, <https://doi.org/10.1007/s10988-018-9249-8>;
- Hom, C. (2008). The semantics of racial epithets. *Journal of Philosophy*, 105, 416-440;
- Hom, C. (2010). Pejoratives. *Philosophy Compass*, 5 (2), 164-185;
- Hom, C. (2012). A puzzle about pejoratives. *Philosophical Studies*, 159, 383-405;
- Hom, C. & May, R. (2013). Moral and Semantic Innocence. *Analytic Philosophy*, 54 (3), 293-313;
- Hom, C. & May, R. (2014). The inconsistency of the identity thesis. *Protosociology*, 31, 113-120;
- Hom, C. & May, R. (2018). Pejoratives as Fiction. In D. Sosa (a c. di) (2018), *Bad words*, Oxford: Oxford University Press, 108-131;
- Horn, L. (1985). Metalinguistic Negation and Pragmatic Ambiguity. *Language*, 61 (1), 121-174;
- Horn, L. (1989). *A Natural History of Negation*. Chicago: University of Chicago Press;
- Jeshion, R. (2013a). Expressivism and the Offensiveness of Slurs. *Philosophical Perspectives*, 27 (1), 231-259;
- Kauppinen, A. (2007). The Rise and the Fall of Experimental Philosophy. *Philosophical Explorations*, 10 (2), 95-118;
- Knobe, J., & Nichols, S. (2013) (a c. di), *Experimental Philosophy*. Oxford, Oxford University Press;
- Liao, S. M. (2008). A defense of intuitions. *Philosophical Studies*, 140 (2), 247-262;
- Machery, E. (2017). *Philosophy within its proper bounds*. London: Oxford University Press;
- Marraffa, M. (2009). Lo studio empirico delle intuizioni, ovvero perché la filosofia ha bisogno della scienza (e in particolare della psicologia sociale). *Sistemi intelligenti*, 2, 317-333;
- Panzeri, F. & Carrus, S. (2016). Slurs and Negation. *Phenomenology and Mind*, 170-180;
- Predelli, S. (2010), From the Expressive to the Derogatory: On the Semantic Role for Non-Truth-Conditional Meaning, in Sawyer, S. (ed.), *New Waves in Philosophy of Language*, Palgrave-MacMillan, New York, 164-185;
- Richard, M. (2008). *When Truth Gives Out*. Oxford: Oxford University Press;
- Schlenker, P. (2007). Expressive presuppositions. *Theoretical Linguistics*, 33 (2), 237-245;
- Swain, S., Alexander, J., & Weinberg, J. M. (2008). The instability of philosophical intuitions: Running hot and cold on truetemp. *Philosophy and Phenomenological Research*, 76(1), 138-155;
- Weinberg, J., Nichols, S. & Stich S. (2001). Normativity and Epistemic Intuitions. *Philosophical Topics*, 29, 429-460;
- Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell.



SECTION

2



# SECTION 2

ETHICS

*Roger Crisp*  
The Methods of Ethics

*Massimo Reichlin*  
On the Idea of a 'Method' in Moral Philosophy

---

ROGER CRISP  
University of Oxford  
roger.crisp@st-annes.ox.ac.uk

---

# THE METHODS OF ETHICS\*

---

## *abstract*

*The paper begins with an account of the intellectual background to Henry Sidgwick's writing of his Methods of Ethics and an analysis of what Sidgwick meant by a 'method'. His broad distinction between three main ethical theories – egoism, consequentialism, and deontology – is elucidated and accepted. Sidgwick's different forms of intuitionism are explained, as are his criteria for testing the 'certainty' of a potentially self-evident belief. Section 3 discusses dogmatic intuitionism (common-sense morality systematized) and Sidgwick's own view, in the light of his requirement for precision in ethics. The final section concerns the implications of Sidgwick's position on disagreement for ethical theory. It is suggested that we have some knowledge in ethics, on which most converge, but not much. The paper concludes with a recommendation for a more irenic and less dogmatic approach to philosophical ethics.*

---

## *keywords*

*Henry Sidgwick, ethical methodology, ethical intuitionism, moral disagreement, dogmatism*

---

\* An earlier version of this paper was presented at a conference on *The Methods of Philosophy*, held at the Università Vita-Salute San Raffaele, Milan, 4-6 October 2017. I thank the organizers for the invitation, and the audience for comments and discussion.

It seems only appropriate, in an issue of a journal dedicated to the the methods of philosophy, and in a paper on the methods of philosophical ethics, to begin with Henry Sidgwick's book *The Methods of Ethics*, first published in 1874 and passing (sometimes rather painfully) through six further editions, until the seventh of 1907, which is now canonical.<sup>1</sup> Of course, my title is not the real reason for beginning with Sidgwick. Rather it is the content of his book, and its quality. C. D. Broad famously said in 1930 that, in his view, Sidgwick's *Methods* is 'the best treatise on moral theory that has even been written', an opinion recently echoed by Derek Parfit in 2011, who suggested that the *Methods* contains 'the largest number of true and important claims' of all books in the history of ethics.<sup>2</sup> How one individuates claims, and judges importance, are of course not straightforward matters, but I think Parfit's suggestion is at least highly plausible. I do not, however, want merely to inflict exegesis of Sidgwick on my readers, though there will be some of that in what follows.<sup>3</sup> I will begin with a little on what Sidgwick himself meant by a 'method', an issue which, despite Sidgwick's reputation for clarity, is in fact rather murky. I will then go on to describe his own philosophical method (in the standard sense) – that is, the way he did ethics, focusing in particular on his intuitionism. I shall then look critically at his intuitionist appraisals of common-sense morality, and at his own proposed alternative, suggesting an alternative to that alternative. I will end with a problem for any normative proposal, and then suggest some implications of that problem for the way we do philosophy now.

### 1. Sidgwick's 'methods'

By the time the *Methods* was published, Sidgwick had for a decade been planning a book that might reconcile 'moral sense', or intuition, with utilitarianism. In the preface to the sixth edition, the editor, E.E. Constance Jones, included some illuminating notes for a lecture by Sidgwick on the development of his ethical views. Although Sidgwick is often, with some justification, described as the third of the great 'classical utilitarians', we can see from his own account of his intellectual journey that, though it began with utilitarianism strictly understood, it soon departed from there and was never to return. Sidgwick tells us that he was initially committed to Millian utilitarianism, which he found

---

<sup>1</sup> Sidgwick, 1907. All otherwise unattributed page references in the text are to this volume.

<sup>2</sup> Broad, 1930, p. 143; Parfit, 2011, p. xxiii.

<sup>3</sup> Many of the arguments below are presented in Crisp, 2006 and Crisp, 2015. Readers may wish to consult those books for further background, elucidation, and development.

liberating in contrast to the ‘arbitrary pressure’ of the dubious, confused, or dogmatic moral rules he had been taught. That commitment sat alongside antipathy to the views of the influential Cambridge philosopher William Whewell, whose *Elements of Morality* (1845) led Sidgwick to the view that ‘intuitional’ ethics was, in comparison with mathematics, ‘hopelessly loose’. (Note, by the way, Sidgwick’s rather confusing usage of ‘intuitionism’ to refer both to the normative, first-order, deontological view he rejected, and to a range of epistemological positions, one of which he accepted. More on that below.)

Sidgwick understood Mill to hold that each person seeks her own happiness (psychological hedonism) and that each person ought to promote the happiness of all (ethical hedonism). He found both views attractive, not yet seeing the potential inconsistency between them. As he came to recognize the possibility of conflict between ‘interest’ and ‘duty’, Sidgwick began to think seriously about egoism as a normative view, and concluded that it is the opposition between ‘interest’ (that is, self-interest) and duty or the general good that is central to ethics, rather than that between ‘intuitions’ and hedonism. This, he says, explains the structure of the *Methods*, in which, after an introductory book, egoism (interest) is discussed in book 2, intuitionism (duty) in book 3, and utilitarianism (the general good) in book 4. Further, Sidgwick decided that, despite his aversion to first-order intuitionist ethics, the only way to ground a utilitarian justification for sacrificing one’s own happiness for the sake of others was through a fundamental ethical intuition: ‘I must somehow see that it was right for me to sacrifice my happiness for the good of the whole of which I am a part’.

Where was he to find such an intuition? Sidgwick returned to Kant, and, though impressed by Kant’s view that what is right for any person must be right for all persons in similar circumstances, felt that this notion of universalizability was insufficiently substantive to ground a principle of duty in opposition to egoism. This led Sidgwick to Joseph Butler, who he believed also accepted a ‘dualism of the practical reason’. It was under the influence of Butler that Sidgwick rejected psychological hedonism and accepted the existence of disinterested, other-regarding motivation. Further, since Butler’s powerful critique of utilitarianism also worried Sidgwick, and Sidgwick realized that he himself was already an ‘intuitionist’ by accepting both the Kantian thesis about rightness and the need for an intuitionist foundation for utilitarianism, he decided to reconsider intuitionist ethics itself. Unsurprisingly, Sidgwick decided to study not Whewell, but Aristotle, seeing his *Nicomachean Ethics* as an impartial attempt to make consistent the common-sense morality of his day. Sidgwick set out to do the same for nineteenth-century British morality, and 3.1-9 was the first component of the *Methods* to be completed. Again, more below on how he got on.

The subject of philosophical ethics, Sidgwick says, is the ‘methods’ of ethics, where a method is ‘any rational procedure by which we determine what individual human beings “ought” – or what it is “right” for them – to do, or to seek to realize by voluntary action’. Such a ‘procedure’ need not be a process. Sidgwick allows that a version of intuitionism according to which we have immediate insight into the rightness of certain actions is proposing a method (Sidgwick, 1907, p. 4). Nor is Sidgwick to be understood as suggesting that only actions matter in ethics, and not, say, the feelings or the characters of agents. Indeed he elsewhere allows that the common-sense conception of virtue includes the emotions (Sidgwick, 1907, pp. 222-3) and that ethics should construct ideals of character (Sidgwick, 1907, p. 393). But discussion of such topics is significant only in so far as it is related to the primary question of ethics – how we should act, that is, what we have overall reason to do. Like Aristotle, Sidgwick sees philosophical ethics as essentially practical (Sidgwick, 1907, p. 1; 4).<sup>4</sup>

---

<sup>4</sup> Aristotle, *Nicomachean Ethics*, 1103b26-9.

The definition of method here leaves it open whether Sidgwick is including ethical theories, which advocate certain basic normative principles, as among the ‘rational procedures’ he has in mind. It appears not. Consider the view that God has implanted in us knowledge of certain apparently non-utilitarian common-sense rules, such as the rule that we should keep promises, because these rules are the best way to promote the utilitarian end of general happiness. According to Sidgwick, this view constitutes a rejection of the *method* of utilitarianism, though not of the utilitarian *principle* (Sidgwick, 1907, p. 85). Ethics, however, as Sidgwick points out, is ‘sometimes considered as an investigation of the true ... rational precepts of Conduct’ (Sidgwick, 1907, pp. 2-3), and he himself implies that we are interested in the principles that determine which conduct is ultimately reasonable (Sidgwick, 1907, pp. 5-6; see 14). In other words, philosophical ethics is an inquiry into what grounds or justifies our actions and any decision-procedure we adopt, and so we might wonder why Sidgwick emphasizes methods rather than ultimate principles (Jerry Schneewind plausibly suggests that Sidgwick was influenced by analogies between scientific and ethical methods.) Sidgwick’s book should perhaps have been titled *The Ultimate Principles of Ethics*, those principles each being a different statement of our ultimate reasons for action; and if it had been so titled some of the unclarities introduced by his focusing on methods might have been avoided.

Sidgwick says that, by adopting some plausible assumption or other, one can connect almost any method with almost any ultimate principle. But he limits his own discussion to those methods ‘logically connected with the different ultimate reasons widely accepted’ (Sidgwick, 1907, p. 78). In other words, he is looking for certain commonly alleged ultimate reasons for action (such as the promotion of happiness), and the decision-making procedure that most naturally rests on each such reason. For example, egoistic hedonism consists in the acceptance of the agent’s own happiness as ultimate reason or end, and its method will be the attempt by the agent to maximize her own happiness. But since method and end can come apart – as in the case of Sidgwick’s own utilitarianism – one might wonder why Sidgwick appears to require that, before any ultimate reason can be given a place in his discussion, its related method must be present in common-sense morality. Imagine that common-sense morality made no room for utilitarian decision-making. The utilitarian principle could still capture what justifies the common-sense method or decision-procedure. Further, his emphasis on method over principle can lead to philosophical distortion. The focus on method may explain, for example, why Sidgwick is so ready to find utilitarianism within common-sense morality, because people sometimes decide what to do by trying to work out what would do the most good from the impartial point of view. This is evidence merely of a principle of beneficence, alongside others, and that is a far cry from utilitarianism as usually understood, according to which this is the only ethical principle.

The distinction between what has become known as the ‘criterion’ of morality – what it is that makes actions right or wrong – and decision-procedures has become clearer since Sidgwick’s time, though there is no doubt that it is at work in some of his arguments. But I do think that his decision to focus on just three ‘methods’ – that is, three broadly ethical theories – egoism, intuitionism (what we might call ‘deontology’), and utilitarianism – is reasonable. These are the three positions which have come to the forefront in the history of western ethics since Socrates, though of course egoism went largely into retreat in the twentieth century, two reasons for that being perhaps G.E. Moore’s somewhat unpersuasive but widely accepted objections along with Victorian distaste for the position (distaste also felt by Sidgwick, who describes the view as ‘despicable’ in earlier editions of the *Methods*, and merely ‘ignoble’ in the sixth and seventh). The debate in normative ethics, then, is between rational or normative egoism, deontology, and utilitarianism (or perhaps more broadly ‘consequentialism’). As I have said, Sidgwick began by elucidating egoism, focusing in particular on its conception

of well-being and arguing for hedonism over ‘perfectionism’. What he says about hedonism is brilliantly insightful, but in fact he could have largely left open the question of what well-being consists in, except in so far as it is claimed to involve virtue – another issue dissected beautifully by him in 2.5. When it comes to ‘dogmatic’ intuitionism – his somewhat questionable name for first-order deontology – his focus is very much more on the meat of the theory. Indeed his two-hundred page discussion of that view in book 3 is almost certainly the best elucidation of that position in the history of philosophy, and very sadly neglected, especially by most of its current adherents, who believe Sidgwick is too boring and too utilitarian to read (in fact, as anyone who reads him seriously will soon realize, he is neither!). Before we get to that, however, let us briefly look at intuitionism in general.

According to Sidgwick, intuition is a doxastic faculty (Sidgwick, 1907, pp. 4-5; 28; 275), nothing more, or less, than a capacity for forming beliefs of a certain kind, with the possibility thereby of acquiring knowledge. These beliefs, because of their apparent source in intuition, may be called ‘intuitions’, and Sidgwick feels himself entitled to assume their *existence*, though not their *truth* (Sidgwick, 1907, pp. 210-12):

## 2. Intuitionism

I wish therefore to say expressly, that by calling any affirmation as to the rightness or wrongness of actions ‘intuitive’, I do not mean to prejudge the question as to its ultimate validity, when philosophically considered: I only mean that its truth is apparently known immediately, and not as the result of reasoning. I admit the possibility that any such ‘intuition’ may turn out to have an element of error, which subsequent reflection and comparison may enable us to correct; just as many apparent perceptions through the organ of vision are found to be partially illusory and misleading: indeed the sequel will show that I hold this to be to an important extent the case with moral intuitions commonly so called. (Sidgwick, 1907, p. 211; see 34n2)

To this extent, then, Sidgwick’s intuitionism is cognitivist. Intuitions for him are not to be understood as, say, certain sentiments or attitudes, but as non-inferential (or ‘immediate’: Sidgwick, 1907, p. 98) rational beliefs which present themselves to the subject as true (and hence as potentially false) -- that is, as apparently self-evident (see e.g. Sidgwick, 1907, p. 338). Each person has to make up her own mind about whether she has any such beliefs, and at this point Sidgwick is characteristically cautious, noting that, in his experience, people seem likely to identify as moral intuitions non-cognitive states such as mere impulses to, or unspecific preferences for, action, or non-intuitive cognitive states, such as conclusions from rapid or semi-conscious processes of reasoning or familiar and unreflectively held opinions. The fact that Sidgwick denies the status of intuition to the latter shows that mere non-inferentiality is insufficient for a belief to be an intuition: it must also present itself as self-evident to the subject. If, as soon as the subject holds her customary belief up to the light it seems not true to her, then that belief was and is not a moral intuition. Some find talk of a faculty of intuition mysterious,<sup>5</sup> and may seek to undermine the idea through some developmental or evolutionary account (Sidgwick, 1907, pp. 212-14). Sidgwick refuses to allow that mere knowledge of the causal history of certain apparent cognitions provides any ground for distrusting them. All of our cognitive faculties have such a history, and hence so do all of our beliefs (including the beliefs of those who criticize intuitions). And the subject-matter of ethical judgements is such that they cannot be inconsistent with

---

<sup>5</sup> See e.g. Mackie, 1977, pp. 38-9.

any merely descriptive psychological or physical claims. Nor, even if our capacity for moral cognition had been present from the start, would that give it any special claim to correctness. None of this is to deny, of course, that some more specific causal account of the origin of certain beliefs might not increase the probability of their being mistaken.

In the nineteenth century, intuitionists were often seen in opposition to so-called *inductivists*. Having presented a broad conception of intuition as ‘immediate [i.e. non-inferential] judgement as to what ought to be done or aimed at’, Sidgwick is careful to situate his intuitionism in the context of contemporary debate (Sidgwick, 1907, pp. 97-8), noting quite correctly that the parties in the debate were commonly talking at cross purposes, since each was claiming to know different things. The inductivists claimed inductive knowledge of the pleasantness of certain actions, whereas intuitionists focused on the rightness (or wrongness) of those actions. Rational or normative hedonism itself cannot be known inductively. It must be either grasped intuitively, or inferred from other premises at least one of which must include a normative intuition.

What we see here is a standard argument for foundationalism in epistemology. Someone inclined towards a non-foundationalist approach, such as some form of coherentism, may well accept Sidgwick’s negative argument against inductivism (that induction alone cannot ground an ethical theory), but deny that foundationalism is the most plausible alternative. Rather, it might be claimed, for example, that hedonism provides the most consistent and coherent fit with other beliefs we have about goodness, rightness, or the world in general. Sidgwick’s own relationship with coherentism is a complex one, as we shall shortly see.

Sidgwick outlines three forms of intuitionism (Sidgwick, 1907, p. 102). The first is ‘perceptual’ intuitionism, according to which each person should do what they think right in each case as it arises. Sidgwick plausibly rejects this view, on the basis that most people find their own particular intuitions open to doubt, non-comprehensive, inconsistent over time, and indeed often in conflict with those of others, and will want to appeal to some general rules or moral theory to support their particular judgements. The second form, as I have mentioned, is ‘dogmatic’ intuitionism. According to the dogmatic intuitionist, general moral rules are implicit in common-sense moral thought, and the task of the philosopher is to elucidate and systematize them as far as possible. Dogmatic intuitionism is not, then, entirely unreflective; indeed it is the result of reflection on the *dogmata* – the beliefs – of common sense. But Sidgwick no doubt felt the name appropriate partly because he believed the view to be *insufficiently* reflective, and hence unable to provide a coherent underpinning for common-sense morality itself, an underpinning which would enable the agent to know exactly which obligations she was under in each situation in which she found herself. What is particularly objectionable to Sidgwick about dogmatic intuitionism, then, is not its starting from common-sense morality, but its readiness to end there also without having removed the indefiniteness which is, he thought, as inappropriate in ethics as in a legal system.

That brings us to Sidgwick’s preferred view: ‘philosophical’ intuitionism. Here Sidgwick is seeking ‘one or more principles more absolutely and undeniably true and evident, from which the current rules might be deduced, either just as they are commonly received or with slight modifications and rectifications’. He goes on:

[W]e conceive it as the aim of a philosopher, as such, to do somewhat more than define and formulate the common moral opinions of mankind. His function is to tell men what they ought to think, rather than what they do think: he is expected to transcend Common Sense in his premises, and is allowed a certain divergence from Common Sense in his conclusions. It is true that the limits of this deviation are firmly, though indefinitely, fixed: the truth of a philosopher’s premises will always be tested by the

acceptability of his conclusions: if in any important point he be found in flagrant conflict with common opinion, his method is likely to be declared invalid. Still, though he is expected to establish and concatenate at least the main part of the commonly accepted moral rules, he is not necessarily bound to take them as the basis on which his own system is constructed. Rather, we should expect that the history of Moral Philosophy – so far at least as those whom we may call orthodox thinkers are concerned – would be a history of attempts to enunciate, in full breadth and clearness, those primary intuitions of Reason, by the scientific application of which the common moral thought of mankind may be at once systematised and corrected. (Sidgwick, 1907, pp. 373-4)

Sidgwick's readiness to assess his own views by the lights of the common-sense morality he is so doubtful about represents, I suggest, an unwillingness entirely to surrender the methodology of Aristotelian dialectic, in which one's own philosophical conclusions must be tested against what is commonly thought. If he had done so, his theory would certainly have been more radical; but it would also have been epistemically purer and hence, I suggest, more systematic. Since Sidgwick nowhere explains exactly how much epistemic weight to give to common-sense morality – that is, how far a philosopher's conclusions may diverge from that morality – he is leaving it open for a dogmatic intuitionist to use the many cases in which utilitarian conclusions are violently at odds with common-sense morality as an internal argument against Sidgwick's own arguments for utilitarianism, as well as unclear how we are to decide between two theories, one of which may have more self-evidence on its side but is less consistent than the other with common-sense morality.

But Sidgwick has good advice for the philosophical intuitionist, pure or impure, on how to proceed. He offers four tests which any allegedly self-evident intuition must pass if it is to be judged of the 'highest certainty' (Sidgwick, 1907, pp. 338-42). The first three are relatively straightforward:

- (I) *Clarity*. 'The terms of the proposition must be clear and precise.'
- (II) *Reflection*. 'The self-evidence of the proposition must be ascertained by careful reflection.'
- (III) *Consistency*. 'The propositions accepted as self-evident must be mutually consistent.'

The fourth is more tricky, and stated indirectly:

- (IV) *Non-dissensus*.

Since it is implied in the very notion of Truth that it is essentially the same for all minds, the denial by another of a proposition that I have affirmed has a tendency to impair my confidence in its validity.... And it will be easily seen that the absence of such disagreement must remain an indispensable negative condition of the certainty of our beliefs. For if I find any of my judgments, intuitive or inferential, in direct conflict with a judgment of some other mind, there must be error somewhere: and if I have no more reason to suspect error in the other mind than in my own, reflective comparison between the two judgments necessarily reduces me temporarily to a state of neutrality. And though the total result in my mind is not exactly suspense of judgment, but an alternation and conflict between positive affirmation by one act of thought and the neutrality that is the result of another, it is obviously something very different from scientific certitude.



We will come back to the fourth condition later. But let me now consider Sidgwick's discussion of dogmatic intuitionism in the light of his three conditions.

### 3. Dogmatic Intuitionism and Sidgwick's Dualism

Towards the end of book 3, after his long survey of common-sense morality, Sidgwick undertakes to bring together the various threads of his argument. He begins by reminding us of his general aim (Sidgwick, 1907, pp. 337-8): to decide whether common sense can provide the basis for 'clear and precise principles commanding universal acceptance', on the understanding that such principles may be arrived at only through philosophical refinement of common-sense morality itself. These principles, in so far as they are self-evident, must of course be non-derivative and ultimate (Sidgwick, 1907, pp. 350-51).

Sidgwick draws a general connection between the clarity and non-dissensus conditions, arguing that attempting to make a common-sense principle precise will lead one into disagreement about it. In other words, even if one becomes clear in one's own mind about some complex principle – such as the principle, say, that one should keep a promise made to a living person, in full knowledge of its risks, to do something morally acceptable, and so on – one will find epistemic peers who disagree with one about that principle once refined. Common-sense moral principles are either too vague to meet the clarity condition, or, if they meet that condition, fail to meet the non-dissensus condition.

But there are problems with Sidgwick's critique. His clarity principle is too demanding: he requires an acceptable ethical principle to be so precise that it can cover any plausible eventuality without the need for any judgement on the part of the person acting on it. Sidgwick holds a view of ethics as aiming at the level of precision appropriate to a natural science, and this explains a good deal of his antipathy towards common-sense morality. Sidgwick ignores or plays down the notion of practical wisdom as a capacity to judge correctly in particular cases, perhaps in the light of general 'prima facie' or pro tanto principles. This conception of principled judgement plays an essential role in many central versions of dogmatic intuitionism, from Aristotle to Whewell (and of course on to Ross and others who wrote after Sidgwick).

Is Sidgwick's own utilitarian principle in a better position than pluralistic intuitionism? Judgement is involved in accepting utilitarianism to start with, but it will also be involved in assessing the implications of the principle in particular cases. Consider some apparently simple case in which I have to decide between two delicious desserts. Sidgwick himself, in his discussion of empirical hedonism in 2.2-3, brings out how difficult such a decision might be. I shall have to recall earlier experiences with each kind of dessert, analyse the quality of each of those on offer, and then make a judgement based on the evidence before me. It is not clear why such judgements are any easier than trying to decide, for example, what form my expression of gratitude to some benefactor should take. Judgement is inescapable. What matters is how best to ensure that one's judgements are correct.

Of course, pluralistic intuitionists should not be against system, and most are not. Consider, for example, Aristotle's discussion of how to rank the different claims on us of those differently related to us, as relatives, citizens, and so on, in *Nicomachean Ethics* 9.2, or Ross's list of prima facie duties. But we have to remember that, however much systematizing we do, the context of human decision-making is sufficiently complex and unpredictable that any plausible ethical theory has to allow some room for individual judgement about particular cases.

What is Sidgwick's own preferred alternative to common-sense morality? Sidgwick accepts:

*Benevolence*: One ought to aim at the universal good.

He does not, however, go for straight utilitarianism, accepting:

*Prudence:* One ought to aim at one's good on the whole.

This combination, of course, is his version of the 'dualism of practical reason', which leads him to despair, as he cannot find a principled solution to problem cases in which my own good and the universal good come apart. But if Sidgwick had recognized the importance of practical judgement, he could have allowed for a more plausible form of the dualism, in which each agent has reason to aim at her own good *and* reason to aim at the universal good, the circumstances she finds herself in determining which reason, in those circumstances, has greater weight.

Now return to W.D. Ross, who says:

If, so far as I can see, I could bring equal amounts of good into being by fulfilling my promise and by helping some one to whom I had made no promise, I should not hesitate to regard the former as my duty... [and] normally promise-keeping ... should come before benevolence.<sup>6</sup>

Ross, then, appears committed to the following principle:

*Promise:* Any agent has ultimate reason to keep promises.

I suspect that many people, perhaps the majority, would accept Promise, and this leaves us with a philosophical stand-off between deontologists on the hand and utilitarians or consequentialists on the other (and, if I am allowed a third hand, rational egoists as well). Remember Sidgwick's fourth condition – non-dissensus – which requires us to suspend judgement on any fundamental ethical proposition if we find that some epistemic peer disagrees (or, even worse, *might* disagree) with it. This condition strikes me as highly credible, and its implications might seem deeply worrying, since I see no plausible way to show that many of those who disagree with me in ethics are not (at least!) my epistemic peers. Is normative philosophical ethics, then, left in complete chaos? Not entirely. First, there is enough consensus on some principles for us even perhaps to say that we know them. Consider, for example:

*Own Pain:* Any agent has a reason to minimize (undeserved) suffering in their own life.

I find it hard to see how any sane person could disagree with this principle. Certainly all the sane people I have met seem to act in accordance with it. But note that it really doesn't take us very far, practically. It is just one pro tanto reason, and cannot be said to be our only reason (unless we should also accept that all of us should commit suicide as soon and as painlessly as possible). Not even the following principle will pass the non-dissensus test:

*Another's Pain.* Any agent has a reason to minimize the suffering of another.

Egoists will disagree with this, and though I think the many egoists in the history of philosophy to be mistaken, I cannot find any way to make them all my epistemic inferiors. Where, then, does this leave philosophical methodology in ethics? In my book *Reasons and*

#### 4. Disagreement and its Implications

---

<sup>6</sup> Ross, 1930, pp. 18-19.

*the Good*, I suggested that there are at least three strategies worth considering: resignation, impartiality, and debate. The first would involve one's refraining from any kind of philosophical debate about normative theory. This strategy, however, seems excessively pessimistic about the prospects for convergence on the truth in ethics and resulting ethical progress. The example of *Own Pain* suggests that truth is available, and there are at least some signs of *apparent* ethical progress, such as the recognition of racism, sexism, and speciesism. So, on the assumption that progress in ethics is possible, one might seek to engage in *impartial* consideration of and debate concerning the various current normative theories. There is more to be said for this strategy than for resignation, and it is certainly likely to be part of the best overall philosophical package. But the fact remains that one will usually persist in one's attachment to the normative principles that strike one as plausible, and one is thus in an especially good position to spell out those principles and display their advantages to others. So this leaves us with the final strategy, in which debate between the advocates of the different normative theories continues. But carried out between those who have suspended judgement as to the correctness or otherwise of the view which they themselves find attractive, such debate would be less adversarial and more constructive than much in philosophy at present. This would have several significant advantages. First, each participant would be more likely to notice the faults in her own position and the advantages in those of others. Second, philosophers would see that there is often greater epistemic benefit in discussing issues with those of radically different views than with some clique of one's own. Third, the aim of debate would be not the victory of one's own position but convergence on some truth, which might well be a conglomeration of various elements from several existing ethical theories. Ethical enquiry must be informed by a spirit of impartiality, in which those who propose normative principles are prepared both to hold up those principles to the light of rational reflection and the arguments of others, actual or imagined, and to look enthusiastically at the views of others, in search of enlightenment rather than dialectical victory. Critical argument, of course, would continue to be the mainstay of moral philosophical discussion, but if it were freed of its unjustified dogmatism there would be a greater likelihood of convergence on the truth.<sup>7</sup> Contemporary philosophy seems to me analogous in some ways to the following. Imagine a group of equally experienced cavers, lost underground in the darkness, and wanting to get out as soon as possible. They share ideas, and it turns out that they all disagree on the best strategy. Only one of them, at best, can be right about the way to go. Now imagine that each decides to spend the time she has available trying hard to persuade her colleagues to agree with her, focusing only on what she sees as the weaknesses of their views and the strengths of her own. If I were lost, I would much prefer to be in a group whose members, though prepared to state their own views as well as possible, were also ready seriously to look for flaws in their own view and advantages in the views of others. Do we, or do we not, seriously want to find our way out of the cave? If so, then we should change the way we currently do moral philosophy.

### REFERENCES

- Broad, C.D. (1930) *Five Types of Ethical Theory*. London: Routledge and Kegan Paul;  
Crisp, R. (2006) *Reasons and the Good*. Oxford: Clarendon Press;

---

<sup>7</sup> In discussion, Timothy Williamson pointed out the advantages of the adversarial approach in certain legal systems. The overall decision in such systems, however, lies not with the adversaries but with an impartial judge. My suggestion is that philosophical adversaries seek to incorporate more of the impartial judicial perspective into their own mode of debate.

- (2015) *The Cosmos of Duty: Henry Sidgwick's Methods of Ethics*. Oxford: Clarendon Press;
- Mackie, J. (1977) *Ethics: Inventing Right and Wrong*. Harmondsworth: Penguin;
- Parfit, D. (2011) *On What Matters*, vol. 1. Oxford: Oxford University Press;
- Ross, W.D. (1930) *The Right and the Good*. Oxford: Clarendon Press;
- Sidgwick, H. (1907) *The Methods of Ethics*. 7<sup>th</sup> edn. London: Macmillan.



---

MASSIMO REICHLIN

Università Vita-Salute San Raffaele  
reichlin.massimo@hsr.it

---

# ON THE IDEA OF A ‘METHOD’ IN MORAL PHILOSOPHY\*

---

## *abstract*

*The paper discusses the two meanings that ‘method’ is often assumed to have in moral philosophy: the epistemic meaning, according to which a method is a procedure to reach moral knowledge, and the normative meaning, according to which it is a criterion of right and wrong in actions. The origin of these two, clearly connected meanings can be traced to Sidgwick’s work The Methods of Ethics. It is argued that Sidgwick’s seminal idea of a “reflective equilibrium” is a valuable and lasting contribution to the debate on moral epistemology; however, Sidgwick’s characterisation of the different normative options is biased against non consequentialist approaches by its concentration on “methods”, rather than on theories and “ultimate reasons”. This consequentialist bias still lingers in contemporary ethics.*

---

## *keywords*

*method, Sidgwick, reflective equilibrium, consequentialism, deontology*

---

\* I wish to thank Roger Crisp for some helpful comments on an earlier draft.

1. *The Methods of Ethics* is the title of one of the most important works in the history of moral philosophy, written by Henry Sidgwick. The influence of this work is such that, when talking about 'method' in moral philosophy, reference is generally made to one of the two things that Sidgwick meant with that word. The two different but connected meanings, that unfortunately Sidgwick fails to clearly distinguish, are the following: On the one hand, 'method' stands for a procedure that must be followed in order to reach moral knowledge, or any surrogate for knowledge that can be offered in an expressivist context (Sugarman and Sulmasy, 2010); on the other hand, it refers to a normative criterion for reaching conclusions as to the rightness or wrongness of human conduct, or to the goodness or badness of some trait of character or state of affairs (Baron, Pettit and Slote, 1997). In the first meaning, the issue of method belongs with moral epistemology, in the second it is part of normative ethics. In this paper I will argue: a) that what Sidgwick says on the epistemic side, in one of its possible interpretations, is basically sound and b) that what he says on the normative side, while important and influent, is open to relevant criticism, both in his identification of the options in the 'normative menu', and in the specific characterisation of the main adversary of what he calls 'the method of utilitarianism'. The limits of Sidgwick's treatment – I will eventually suggest – have had a huge impact on contemporary normative discussion.
2. Sidgwick famously distinguished three methods of ethics: egoistic hedonism, intuitionism, and universalistic hedonism or utilitarianism. These three – to which, according to Sidgwick, all other methods are reducible – are the names of three normative theories, or at least can be so conceived and are often in fact conceived. However, as already mentioned, the notion of a method also has an epistemic significance. In fact, following earlier representatives of the utilitarian tradition, such as John Stuart Mill, Sidgwick presents the two hedonistic methods as based on the empirical or a posteriori strategy of establishing specific moral conclusions by calculating consequences in terms of the pleasure produced, and the intuitional method as based on the a priori strategy of grasping fundamental normative axioms and deriving from them particular moral conclusions. In Sidgwick's use, the two meanings are interrelated in a complex way, because the main result of his book is to show the substantial convergence or conciliation of intuitionism and utilitarianism, both at a normative and at an epistemic level. On the one hand, while 'dogmatic intuitionism' cannot receive a philosophical vindication, utilitarianism offers a philosophical systematization of the morality of common sense, inasmuch as the latter is "unconsciously

utilitarian”; on the other hand, Sidgwick’s critique of the empiricist epistemology of earlier utilitarians leads him to ground utilitarianism in the really self-evident first principles he establishes, and thus to present it as nothing but “the final form into which intuitionism tends to pass, when the demand for really self-evident first principles is rigorously pressed” (Sidgwick, 1907, p. 388). In short, the upshot of Sidgwick’s work is the convergence of utilitarianism and *philosophical* intuitionism. When in 1930 Charles Broad first noted the ambiguity of the term ‘intuitionism’ in Sidgwick’s treatment (Broad, 1930, p. 206), and proposed to use Bentham’s word ‘deontology’ to denote the normative system based on a priori principles, he was in a sense making a perfectly appropriate point; however, he was also somehow obscuring the intrinsic connection between the two elements which is a relevant feature of Sidgwick’s view. Intuitionism is the normative view accepting standards of right and wrong that are based on a priori knowledge – and in the end, for all the remaining differences, the same holds for utilitarianism as well<sup>1</sup>.

I will be comparatively brief on the epistemic aspect of the discussion. The nature of Sidgwick’s epistemic approach has been the object of a complex discussion, and I can’t get very much into this scholarly debate in this context (Singer, 1974; Deigh, 2007; Skelton, 2010). However, it seems to me that Sidgwick was basically right in rejecting the empiricist method of his utilitarian predecessors, and in showing that normative thinking cannot proceed without making appeal to intuition at some point; for it is clear that fundamental ethical concepts, such as ‘right’ or ‘ought’, cannot be reduced to any physical or psychological facts, such as desiring or preferring. As Sidgwick says, moral judgments “cannot legitimately be interpreted as judgments respecting the present or future existence of human feelings or any facts of the sensible world” (Sidgwick, 1907, p. 25). Moreover, Sidgwick was also unconditionally right in claiming that moral philosophy must take its start from reviewing the ordinary moral convictions of humanity, and then move forward by attempting to extract from them some principles or axioms that can help remove inconsistencies and systematize the sometimes chaotic beliefs of common sense. In fact, it is far more likely that we stumble into some true moral principle by starting from our ordinary experience of what it is to be under a moral obligation than by applying our rational tools to moral thinking conceived of in a practical vacuum.

3.

Of course, Sidgwick declared that moral philosophy must do more than simply reformulate common moral beliefs: it must transcend common sense in order to tell human beings what they ought to think, instead of merely clarifying what they already think. However, a) Sidgwick clearly grants some initial authority to the rules of common sense. In fact, he views the “current rules” as constraining the work of philosophical intuitionism, since the self-evident moral principles must be such that those rules can be deduced from them, “either just as they are commonly received or with slight modifications and rectifications” (Sidgwick, 1907, p. 102); b) he accepts that “the truth of a philosopher’s premises will always be tested by the acceptability of his conclusions: if in any important point he be found in flagrant conflict with common opinion, his method is likely to be declared invalid” (Sidgwick, 1907, p. 373); c) he declares that his fundamental axioms of Justice, Prudence and Benevolence are themselves part of that morality of common sense, along with other principles whose self-evidence is illusory; d) he declares that the fact that it sustains the general validity of current moral

---

1 At one point Sidgwick calls the normative view that excludes the consideration of consequences “the narrower sense” of ‘intuitional’, distinguishing it from the “the wider sense”, that indicates the epistemic view affirming self-evident normative principles (Sidgwick, 1996, p. 102).



judgments is part of the “proof” of utilitarianism.

The method here presupposed is, therefore, a complex blending of foundationalist and coherentist elements (Brink, 1994; Phillips, 2011, pp. 65-76): philosophical ethics takes its start from common morality, transcends it by establishing the axioms that correct and systematize the rules of common sense, but then, in the last resort, is tested against common beliefs; “a certain divergence from Common Sense” (Sidgwick, 1907, p. 373) is allowed in the conclusions, but the subversion of ordinary moral thinking would be epistemically suspect. This method does show some similarity with contemporary proposals of “wide reflective equilibrium” (Daniels, 1979; Scanlon, 2014), even though the discussion is open on this point. In any case, while moral philosophy cannot simply reiterate and perhaps systematize what may in fact turn out to be our sheer prejudices, it cannot aim at rebuilding our moral system *ex novo*, either on the basis of conceptual or linguistic intuitions concerning the notions of morality or rationality. Some commentators have claimed that Sidgwick should be regarded as “a modest [*i.e.* not dogmatic] foundationalist” (de Lazari-Radek and Singer, 2014, p. 111), and that this position is not far from those conceptions of wide reflective equilibrium, such as Daniels’ or Scanlon’s, that allow for the (fallible) recognition of objective moral truths, and do not confine the goal of the moral enterprise to the search for intersubjectively constructed moral principles. This may be right, but still, to say that, at the end of the inquiry, “we will be able to demonstrate that none, or virtually none, of our existing moral judgments are credible” (de Lazari-Radek and Singer, 2014, p. 113) misinterprets Sidgwick’s epistemic approach, and strikes me as highly implausible and perhaps impossible. It is one thing to say that no single judgment is excluded from possible revision; quite another to say that *all* our present beliefs may be wrong or false. In fact, we construct our moral judgments, such as the one that we ought to maximise overall welfare or to give to each his or her due, starting from ordinary beliefs, such as that pleasure is good or that each person has equal worth. We can surely test any such ordinary belief against others, and perhaps discard some of them as inconsistent with our whole body of beliefs; however, should we declare all our ordinary beliefs incredible, we would simply lack any place where to start from in order to construct the judgments and principles of our moral theory. In so far as his moral epistemology was hybrid between foundationalism and coherentism, I assume that this is also what Sidgwick would say, and, therefore, that he would dismiss the attempt to reject all our current judgments and “replace them with the judgments that follow from the [soundest] moral theory” (de Lazari-Radek and Singer, 2014, p. 113).

An a posteriori confirmation of this claim is the fact that attempts to ground straightforwardly foundationalist systems of morality bracketing all substantive beliefs, such as those embarked on by Richard Hare or Richard Brandt, have been shown to presuppose at some point substantive (and possibly non universally valid) moral intuitions<sup>2</sup>; in general, any ‘foundationalist’ definition of the space of morals, presuming to bring out criteria of normative judgment without presupposing any particular moral beliefs, inevitably smuggles in some moral content, which is taken for granted as a substitute of full-blown moral intuitions. So, as far as the epistemic dimension of ‘method’ is concerned, it seems to me that there is nothing better that we can do than, in Sidgwick’s wake, to embark on a phenomenology of our moral experience, and try to extract from it some more general principles which we can use to partly correct or refine our beliefs with a view to making them more consistent and systematically connected. This is not to say, of course, that Sidgwick was right, either in precisely stating the really self-evident axioms of ethics, or in the consequentialist-friendly interpretation he

---

2 For comments on Hare’s project, in particular, see Mackie, 1977.

gave of them; nor does this mean that he was charitable and fair in his critique of attempts to systematise the rules of common sense morality within a non-consequentialist framework (Shaver, 2014; Hurka, 2014; Phillips, 2011, p. 100-3). But this is already part of the normative issue.

Sidgwick talked of egoism, intuitionism and utilitarianism as different ‘methods’ of ethics; and it could be objected that talk of different ethical ‘theories’ might have been much clearer and more appropriate. However, contrary to what it is sometimes implied, the three methods discussed by Sidgwick are *not* ethical theories. As a matter of fact, Sidgwick himself defines a method as “any rational procedure by which we determine what individual human beings ‘ought’ – or what it is ‘right’ for them – to do, or to seek to realise by voluntary action” (Sidgwick, 1907, p. 1). This is not exactly the same as a philosophical theory, for, in Sidgwick’s use, the latter would involve systematically justifying one procedure as *the* one and only right method, or demonstrating the final harmonization of the different methods, and providing a complete solution of the chief ethical difficulties and controversies (Sidgwick, 1907, p. 13). Moreover, in Sidgwick’s view, an ethical theory does also encompass the establishment of ethical first principles, that is, a statement and justification of the ultimate reason for action. He clearly distinguishes method from principle (Sidgwick, 1907, p. 85), when he writes that the view requiring to follow the non-utilitarian rules of common sense in order to promote general happiness rejects the *method* of utilitarianism, but not its *principle*; and he clearly says that it is not his “primary aim to establish such principles” (Sidgwick, 1907, p. 14), but to focus on the methods used by common sense (Schneewind, 1977, pp. 191-3).

4.

Sidgwick’s basic purpose – at least in his official declarations – is therefore not to establish or defend an ethical theory. His declared aim is only to diminish the confusion that he finds apparent in ordinary ethical thinking, which uses a variety of methods, without being aware of their partial clash or of the fact that inconsistent principles are used at different times. He just wants to “expound as clearly and as fully as my limits will allow the different methods of ethics that I find implicit in our common moral reasoning; to point out their mutual relations, and, where they seem to conflict, to define the issue as much as possible” (Sidgwick, 1907, p. 14). In particular, his main goal is to critically discuss the morality of common sense, and to reject “the attempt to elevate it into a system of Intuitional Ethics” (Sidgwick, 1907, p. 361); he does not want to provide a systematic foundation or complete discussion of utilitarian ethical theory, following his initial declaration according to which his volume “contains neither the exposition of a system, nor a natural or critical history of systems” (Sidgwick, 1907, p. 12; cf. 78). From this point of view, Sidgwick’s attitude is modest and cautious; nonetheless, it is also ambitious in aiming at the identification of *the* fundamental alternatives among which humans *must* choose when trying to establish the reasonableness of their conduct and to reach a systematic account of morality. And we must investigate whether the alternatives are appropriately defined and adequately reflect ordinary moral experience.

My main claim, as far as the normative side of the discussion is concerned, is that the way in which Sidgwick treats the relationship between methods and principles has several unhappy consequences. First, it leads him to unduly limit the options in the ‘ethical menu’; second, it limits his exploration of non-consequentialist views to the systems of rational intuitionists, such as Whewell – and even in this case, without appropriately rendering justice to them<sup>3</sup>; third, it implicitly reduces all other plausible rivals of utilitarianism to this pattern. As a

5.

---

3 On Sidgwick’s uncharitable treatment of Whewell, see Cremaschi, 2008 and 2011.

consequence, the legacy that *The Methods* left to the twentieth-century ethical debate is a blunt and unsatisfactory alternative between a seriously undermined form of pluralistic rational intuitionism and an admittedly sophisticated form of utilitarianism; an alternative from which contemporary ethical theory still struggles to be released. Roger Crisp is right, in other words, in claiming that “his emphasis on method over principle can lead to philosophical distortion” and that Sidgwick’s book would have gained clarity by centering on *The Ultimate Principles of Ethics* (Crisp, 2015a, pp. 21-2).

The fundamental move that Sidgwick makes in *Methods*, I.1.4 is to derive his three methods from the review of what he considers the two ultimate ends that are “strongly and widely supported” (Sidgwick, 1907, p. 9) by the common sense of mankind, namely, happiness and perfection: the first generates the two hedonistic methods, while the latter is connected to the intuitional method as “a special form” of it. It is surprising, in an otherwise scrupulous writer as Sidgwick, to see that this connection is pretty much assumed without argument<sup>4</sup>: from the few lines that he devotes to this point, we get the idea that respecting the intuitional duties of common sense morality is the means to reach the complete realisation of virtue, which in turn is “the most valuable element of human Excellence” (Sidgwick, 1907, p. 11). As noted by two contemporary admirers, this seems to mean that “In obeying these rules, we achieve the goal of our own moral perfection or excellence” (de Lazari-Radek and Singer, 2014, p. 16). This interpretation is confirmed by the following observation: “If a man accepts any end as ultimate and paramount, he accepts implicitly as his ‘method of ethics’ whatever process of reasoning enables him to determine the actions most conducive to this end” (Sidgwick, 1907, p. 8). The connection so established between the respect of intuitional rules and the end of perfection, therefore, seems to imply that respecting such rules is rational so long as it produces the desired end of moral virtue. In this way, however, a teleological account is presupposed, and an implicit stress on maximisation is also introduced: in fact, according to Sidgwick, the “natural methods rationalised” (Sidgwick, 1907, p. 12) are those indicating the actions *most* conducive to a certain end.

A few comments are here in order. For one thing, it is not clear that respecting those duties is sufficient, or sometimes even strictly necessary, to achieve human perfection. In fact, it is part of common sense to say that human perfection sometimes requires supererogatory action, which goes beyond the accomplishment of standard duties. In the second place, it is also unclear that morality is in itself sufficient for human perfection, which may include intellectual, artistic and other capacities well beyond rule-abidance. Thirdly, linking the intuitional method to the end of perfection does not seem to do justice to the distinction, accepted by Sidgwick himself (Sidgwick, 1907, I.1.2.), between the views of ethics as the search for the rational precepts of conduct, and as the inquiry into the nature of the ultimate good of human life. The first account presupposes that the rightness of conduct does not necessarily depend on its conducing to some end, for it can also be inherent in the action itself: for example, talking in contemporary terms, some conduct may be made inherently right by the fact that it expresses acknowledgment of some fundamental right of other people, even though it does not foster anyone’s perfection, or other desired end. It is a fact, however, that, starting from its very first presentation, the “intuitional method” is always linked by Sidgwick to the search for a goal, albeit one different from happiness. Fourthly, establishing the connection in this way also seems to exclude the possibility to connect the ultimate goal

---

<sup>4</sup> This perhaps being an example of what Thomas Hurka has called a general truth, i.e., that “the more important a topic is, the less time Sidgwick spent on it, and what he said about it was then often far from clear” (Hurka, 2014, p. 151).

of Perfection with a different method, for example the one based on the exemplary power of embodied virtue. This has the unhappy consequence of excluding from consideration Aristotelian virtue ethics and of juxtaposing it to Clarke's or Price's strictly 'deontological' systems as members of the same 'philosophical family'.

On the whole, the notion of a 'method', as presented in Sidgwick's work, seems compromised from the start with a teleological view, according to which the general feature of a normative criterion is to foster the accomplishment of some specified goal. Sidgwick does not accept that 'right' means 'conducive to a desired end' (Sidgwick, 1907, pp. 25-7) – as Moore will do in the *Principia* – but he sees moral value as a goal that can be produced, and therefore equates 'rightness' and 'rational conduct' with conduciveness to a rationally approved end. Since the rules of the "intuitional method" are basically conceived as means to the end of perfection, this method is not understood as having a quite different point from the teleological one, but as a sort of partial deviation from the archetype.

This conclusion may seem too hasty, for at *Methods*, I.6.1 Sidgwick speaks of three different "ultimate reasons for acting or abstaining" (Sidgwick, 1907, p. 78): happiness, excellence or perfection, and duty. He declares that duty, not perfection, is the ultimate reason to which the intuitional rules refer; moreover, later in that chapter, he also observes that "almost any method may be connected with almost any ultimate reason by means of some – often plausible – assumption" (Sidgwick, 1907, p. 83). Therefore, Sidgwick's view seems to be that the intuitional method can be linked to two different ultimate goods, namely, perfection and the accomplishment of duty. And assuming, as he does, the priority of method over ultimate reason justifies his treating perfectionism as a variety of intuitionism. This would show both that Sidgwick is not presupposing a teleological account of rightness and that he was in fact justified in not acknowledging perfectionism as a distinct account.

6.

There are reasons, however, to be sceptical of this conclusion. The first reason concerns perfectionism. While it is true that following intuitional rules of unconditional duty can be a way to pursue moral virtue, it seems that perfectionism is better conceived of as a non-hedonist and non-consequentialist form of teleology: that is, as a form of moral thinking affirming the centrality of expressing certain attitudes, as typically displayed by virtuous agents<sup>5</sup>. This is not necessarily equivalent to respecting unconditional rules of morality. In fact, according to this theory, acting rightly has more to do with expressing certain traits of character in our relationships with others than with observing rules – which also means that very often there will be more than one possibility of acting rightly, contrary to what Sidgwick assumes. Moreover, in contrast to a strictly deontological moral outlook, theories based on character assume that being a virtuous agent is also, not irrelevantly, to know when a rule should be broken or suspended. Finally, as already mentioned, these theories insist that seeking perfection can often involve going beyond the call of duty. A second reason for scepticism concerns the introduction of duty as an ultimate reason parallel to happiness and perfection, a move that seems insufficient to account for the reasonableness of the intuitional method. For one thing, this strategy presents the notion of duty as self-standing, that is, as something in itself justifying: and this claim is difficult to justify outside a specifically Kantian framework. Moreover, assuming duty as the ultimate ground of action is too unspecific to characterise the intuitional method: also in the utilitarian system, in fact, we can fairly say that doing what maximises happiness is a duty. However, in this case duty

---

<sup>5</sup> I am not convinced, on this point, by Crisp's arguments as to the reducibility of virtue ethics to non-principle-based deontology, as far as the notion of right and wrong action is concerned: see Crisp, 2015b.

can be further explained with reference to the idea of happiness, just as, in a perfectionist view, the promotion of perfection can be offered as the ultimate reason for accomplishing duties. However, in the case of intuitionism, reference to duty as the ultimate reason of the intuitional method leads to the uninspiring and almost tautological view according to which, as commented by Crisp “the reason I should keep my promise [...] is that promises should be kept” (Crisp, 2015a, p. 26). So, even though the intuitional method does not treat moral rules as means to the end of moral virtue, the ultimate reason why we should follow those rules is in no way truly explained.

This strengthens the conclusion according to which a major shortcoming of Sidgwick’s enterprise is to have limited the discussion to the methods, without extending it to ethical principles and theories; for to rationalise conduct is not only to provide a method for reaching practical conclusions, but also to provide an ultimate reason to follow that method. Discussing the intuitional systems, Sidgwick does not provide any such reason, but the mere idea of duty; this is why inscribing the rules of the morality of common sense into the utilitarian framework, with its insistence on the ultimate good of happiness, seems to provide the needed rationale for following them (so long as they foster the accomplishment of that end). But Sidgwick fails to see that the intuitional method can also be inscribed in a different overall scheme, providing a different ultimate reason to follow the ordinary rules. And the search for this alternative rationale for what was later termed a ‘deontological’ view of ethics has often eluded ensuing commentators, so that the ordinary rules of morality have appeared, in Sidgwick’s words, as “an accidental aggregate of precepts, which stands in need of some rational synthesis” (Sidgwick, 1907, p. 102), or, in McNaughton’s phrase, as “an unconnected heap of duties” (McNaughton, 1996).

7. The influence of Sidgwick’s book on twentieth-century ethics has been much deeper than that of any other writer, including Mill and Moore. But his insistence on methods over principles, and his failure to explore several other possibilities in ethical theory have been damaging for the ensuing discussion. Leaving aside egoism, Sidgwick’s dichotomy between the two methods – a dichotomy, of course, which he presumed to have resolved – was later on accepted as *the* basic dichotomy between ethical theories; and the whole point of normative ethics has often been whether to accept some form of utilitarianism or some version of rational intuitionism, most notably in Ross’ version. It was in the context of discussing Sidgwick’s view, in fact, that Charles Broad introduced the notion of deontology to characterise the normative element of the ‘intuitional method’. And, as it has been noted by several commentators (Louden, 1996; Shaver, 2011; Sørensen, 2008; Timmermann, 2015), Sidgwick’s treatment of deontology as a sort of deviation from the ordinary ‘teleological’ conception – a deviation with an inbuilt slight look of irrationality – exerted a relevant influence in the formation of the notion. Suffice it to recall that, in introducing the distinction between teleology and deontology, Broad writes that, while the teleologist deems arbitrary to exclude any part of its foreseen consequences, the deontologist assumes that the rightness or wrongness of an action is “completely determined by certain characteristics of a certain restricted part of its total intended consequences” (Broad, 1930, p. 212). So, what is specific about deontology turns out to be the (somewhat arbitrary) decision to exclude from consideration the many more remote consequences that an act may have, and to concentrate on its immediate consequences, indeed on that restricted part of its immediate consequences which bear on the definition of the act: thus, according to Broad, deontology “defines the kind of action under consideration by *one* or *a few* characteristics of its *immediate* consequences; and it claims to see that these *suffice* to make all such actions right (or wrong), and that the more remote consequences and the other characteristics of the consequences will always be irrelevant to the rightness (or wrongness)

of the action” (Broad, 1930, p. 214)<sup>6</sup>.

It is apparent that Broad’s notion of deontology is styled on Sidgwick’s characterisation of the intuitional method as a departure from the ‘basic’ teleological account. And Broad even adds that, while a monistic deontology is a theoretical possibility, in fact deontological theories have tended to be pluralistic; the suggestion being, not even implicitly, that deontological theories lack any principle of unification of their several duties, and that this is a reason of weakness, as compared to teleological ones. In fact, “No one has produced a plausible monistic deontological theory; whilst universalistic hedonism is a fairly plausible form of monistic teleological theory. And this fact has often made people prefer teleological theories, since monism in such matters is more satisfactory to the intellect than pluralism” (Broad, 1930, p. 215).

Sidgwick’s discussion of moral epistemology has provided twentieth-century ethics with a viable and fruitful method for reflecting on morals and gaining moral knowledge. But the translation of Sidgwick’s discussion on ‘methods’ into the present discussion on ethical theories has caused an overemphasis on judgments of rightness and wrongness in contemporary normative theory, and the standardisation of a conception of right action as ‘the best, or most reasonable, thing to do’ or ‘the act achieving the most possible value’. This conception obviously leads to the so called ‘paradox of deontology’ and biases the discussion in favour of consequentialism.

8.

In order to retrieve a plausible view of deontology – or any other label one wishes to attach to a normative theory departing from consequentialism – it is vital not to concentrate on ‘methods’, but on theories, and particularly on what Sidgwick called “ultimate reasons” for moral action. If ethics mainly has to do with causing effects in the world, trying to make it a better place, than consequentialism is the most plausible candidate to the role of the best theory. A sensible non-consequentialist view must start from a different general account, according to which ethics has to do with committing oneself to reflectively endorsed principles, through which we define our relationships with others. In such an account, the notions of moral integrity, self-respect and respect for others will be prominent, and the focus will not be on the method for reaching specific normative conclusions, but on the cultivation of a character based on a continuous review of one’s intentions and reasons for action and a consistent application of appropriate moral principles. The accomplishment of this task, however, is, for a large part, still waiting to be carried out.

## REFERENCES

- Baron, M.W., Pettit, P., & Slote, M. (1997). *Three Methods of Ethics*. Melden: Blackwell;  
 Brink, D. (1994). Common Sense and First Principles in Sidgwick’s *Methods*. *Social Philosophy & Policy*, 11, 179-201;  
 Broad, C. D. (1930). *Five Types of Ethical Theory*. London: Routledge & Kegan Paul;  
 Cremaschi, S. (2008). “Nothing to invite or to reward a separate examination”: Sidgwick and Whewell. *Etica & Politica*, 10, n. 2, 137-181 ([http://www2.units.it/~etica/2008\\_2/CREMASCHI.pdf](http://www2.units.it/~etica/2008_2/CREMASCHI.pdf));

---

<sup>6</sup> Something very similar was in fact already present in Sidgwick’s definition of intuitionism as the view that “certain kinds of actions are right or reasonable in themselves, apart from their consequences; – or rather with a merely partial consideration of consequences, from which other consequences admitted to be possibly good or bad are definitely excluded” (Sidgwick, 1907, p. 200; cf. 337).

- Cremaschi, S. (2011). As Boys Pursue the Rainbow: Whewell's Independent Morality vs Sidgwick's Dogmatic Intuitionism. In P. Bucolo, R. Crisp, & B. Schultz (eds.), *Henry Sidgwick. Ethics, Psychics, Politics* (146-234), Catania: C.U.E.C.M.;
- Crisp, R. (2015a). *The Cosmos of Duty: Henry Sidgwick's Methods of Ethics*. Oxford: Clarendon Press;
- Crisp, R. (2015b). A Third Method of Ethics?, *Philosophy and Phenomenological Research*, 90, 257-273;
- Daniels, N. (1979). Wide Reflective Equilibrium and Theory Acceptance in Ethics. *Journal of Philosophy*, 76, 256-282;
- Deigh, J. (2007). Sidgwick's Epistemology, *Utilitas*, 19, 435-446;
- de Lazari-Radek, K., & Singer, P. (2014). *The Point of View of the Universe: Sidgwick and Contemporary Ethics*. Oxford: Oxford University Press;
- Hurka, T. (2014). Sidgwick on Consequentialism and Deontology: A Critique. *Utilitas*, 26, 129-152;
- Louden, R.B. (1996). Toward a Genealogy of 'Deontology'. *Journal of the History of Philosophy*, 34, 571-592;
- Mackie, J.L. (1977). *Ethics: Inventing Right and Wrong*. New York: Penguin;
- McNaughton, D. (1996). An Unconnected Heap of Duties?. *Philosophical Quarterly*, 46, 433-447;
- Phillips, D. (2011). *Sidgwickian Ethics*. Oxford: Oxford University Press;
- Scanlon, T. (2014). *Being Realistic About Reasons*. Oxford: Oxford University Press;
- Schneewind, J.B. (1977). *Sidgwick's Ethics and Victorian Moral Philosophy*. Cambridge: Cambridge University Press;
- Shaver, R. (2011). The Birth of Deontology. In T. Hurka (Ed.), *Underivative Duty: British Moral Philosophers from Sidgwick to Ewing* (pp. 126-145). Oxford: Oxford University Press;
- Shaver, R. (2014). Sidgwick's Axioms and Consequentialism. *The Philosophical Review*, 123, 173-204;
- Sidgwick, H. (1907). *The Methods of Ethics*. London: Macmillan (repr. Bristol: Thoemmes Press, 1996);
- Singer, P. (1974). Sidgwick and Reflective Equilibrium. *The Monist*, 58, 490-517;
- Skelton, A. (2010). Henry Sidgwick's Moral Epistemology. *Journal of the History of Philosophy*, 48, 491-519;
- Sørensen, A. (2008). Deontology: Born and Kept in Servitude by Utilitarianism. *Danish Yearbook of Philosophy*, 43, 69-95;
- Sugarman J., & Sulmasy D., Eds. (2010). *Methods in Medical Ethics*. Washington D.C.: Georgetown University Press;
- Timmermann, J. (2015). What's Wrong with 'Deontology'?. *Proceedings of the Aristotelian Society*, 115, 75-92.

SECTION

3



# SECTION 3

## PHENOMENOLOGY

*Dermot Moran*

What is the Phenomenological Approach? Revisiting Intentional Explication

*Sara Heinämaa*

Two Ways of Understanding Persons: A Husserlian Distinction

*Roberta Lanfredini*

Phenomenological Empiricism

*Roberta De Monticelli*

The Paradox of Axiology. A Phenomenological Approach to Value Theory

---

DERMOT MORAN

Boston College & University College Dublin  
dermot.moran@bc.edu

---

# WHAT IS THE PHENOMENOLOGICAL APPROACH? REVISITING INTENTIONAL EXPLICATION

---

## *abstract*

*In this paper I outline the main features of the phenomenological approach, focusing on the central themes of intentionality, embodiment, empathy, intersubjectivity, sociality and the life-world. I argue that phenomenology is primarily a philosophy of intentional explication that identifies the a priori, structural correlations between subjectivity and all forms of constituted objectivities apprehended in their horizontal contexts. Intentional description reveals the structurally necessary, meaning-informing interactions between embodied subjectivity (already caught in the nexus of intersubjectivity) in the context of embeddedness in the temporal, historical, and cultural life-world. I shall defend phenomenology as a holistic approach that rightfully defends the role of subjectivity in the constitution of objectivity and recognizes the inherent limitations of all forms of naturalism, objectivism and scientism.*

---

## *keywords*

*phenomenology, description, intuition, intentionality, life-world, correlation*

### **1. Introduction: The Reach of Phenomenology**

Phenomenology is arguably the most influential and most widespread (in terms of geographical reach, especially in the non-anglophone world) philosophical movement of the twentieth century; and, as a method or approach, it continues to be extremely influential in the twenty-first century, especially in philosophy of mind and the cognitive sciences, in the medical humanities (especially philosophy of psychiatry), and in intercultural philosophy, and any discussions involving intentionality, consciousness, embodiment, affectivity, mood, empathy, social relations, and embeddedness in cultural tradition. Phenomenology is a flexible approach and there is not one universally accepted method, nevertheless it has distinctive features which I shall try to outline in this paper.

Phenomenology first emerged as the term for a scientific discipline, initially in the writings of Lambert, Kant and Hegel (Moran, 2000), but then was established as a specific method of doing philosophy by Franz Brentano in his *Descriptive Psychology* lectures (Brentano, 1995a) and by Edmund Husserl in his *Logical Investigations* (Husserl, 2001). In Germany, in the nineteen twenties, phenomenology interacted vigorously with contemporary competing philosophical movements, including the Vienna Circle, German Neo-Kantianism, Life-philosophy (Simmel), Hermeneutics (Dilthey), with the nascent Critical Theory of the first Frankfurt School (Horkheimer, Adorno), and also with Marxism (Arendt, Marcuse, Tran Duc Thao). Later phenomenology was brought into dialogue with Neo-Hegelianism (Kojève, Hyppolite, Sartre, Merleau-Ponty), structuralism and poststructuralism (Foucault, Derrida), and with philosophy of language (Reinach pioneered speech acts, Mulligan, 1987), e.g. the work of J. L. Austin and John Searle (Searle, 1983).

Phenomenology explores its main themes – intentionality, intuition, eidetic insight (*Wesensschau*), perception, embodiment, empathy, intersubjectivity, sociality, historicity, and the life-world – in an original and, judging by the amount of attention given to classical phenomenology by contemporary analytic philosophers, still relevant manner (Gallagher & Zahavi, 2012). These themes have largely been recovered and revisited in contemporary philosophy of mind (Dahlstrom 2015), philosophy of action, philosophy of the cognitive sciences, among other areas (Dennett, 1987; Noë, 2005; Hopp, 2011). Furthermore, phenomenology is currently valued not just for its original and still relevant discussions of the first-person perspective, embodiment, emotions and moods, intersubjectivity, historicity, etc., but also for its trenchant critique of various one-sided standpoints that have dominated much contemporary philosophy, e.g., naturalism (Zahavi 2004, 2010, 2013; Moran 2008, 2013a), objectivism, and scientism (Georgi, 2009). The phenomenological tradition is currently

in the process of being re-invigorated through its engagement with the complex issues of intercultural hermeneutics (Waldenfels, 2007; Lau, 2016).

Both as a strict method – Husserl’s “philosophy as rigorous science” (*Philosophie als strenge Wissenschaft*, Husserl, 2002) and, more generally, as a theoretical approach, phenomenology is now well established not only within theoretical philosophy but also in various forms of applied philosophy and in disciplines, such as sociology (Alfred Schütz, Harold Garfinkel), psychology (Aron Gurwitsch), neurology (Francisco Varela, see Petitot & Varela 1999), psychiatry and psychotherapy (Karl Jaspers, Ludwig Binswanger, Erwin Straus (1966), Amadeo Giorgi (1997), Eugene T. Gendlin, Thomas Fuchs, Louis A. Sass, see Durt et al., 2017), and, more recently, as a method of qualitative analysis in the social and health sciences (Jonathan A. Smith et al., 2009) and in the medical humanities.

Phenomenology, first and foremost, as it evolved throughout the twentieth century and into the twenty-first, is best understood as a *tradition* and an *outlook* or *approach* rather than as a strict *method*. Despite Husserl’s best efforts to stipulate its methodological rigor (Husserl, 2014), each classical phenomenologist (e.g. Scheler, Heidegger, Stein, Sartre, Merleau-Ponty) has developed his or her own distinctive approach. As an established philosophical tradition, phenomenology has bequeathed a rich corpus of texts that are now philosophical classics, e.g. Husserl’s *Logical Investigations* (Husserl, 2001) Heidegger’s *Being and Time* (Heidegger, 2012), Sartre’s *Being and Nothingness* (Sartre, 1979), Merleau-Ponty’s *Phenomenology of Perception* (Merleau-Ponty, 2012), to name but a few, that are currently being rediscovered, especially in the Anglo-American tradition (Szanto, 2012), by philosophers, such as: Hubert L. Dreyfus (1991), Charles Taylor, Sean Dorrance Kelly (2000), Alva Noë, Shaun Gallagher, Evan Thompson (2007), John McDowell, Robert Brandom, among many others.

Phenomenology is also – and this distinguishes it from contemporary philosophy of mind--consciously reflective about what *traditionality*, or ‘life in tradition’, involves. Hence Husserl, Heidegger, Gadamer, Ricoeur, Waldenfels, and others, have played close attention to how subjective and cultural life are interwoven in a layered and sedimented manner, how traditions are handed on, distorted, emptied out of significance and reborn in the ebb and flow of collective social and historical intentionality. It is not just that all humans live in history, and tell their historical narratives, their orientation to existence is such as to be intrinsically historical. Some phenomenologists have been drawn toward Hegel’s or Dilthey’s discussions of historical evolution and development, but phenomenology tends to look more at how humans live temporally in an historical tradition. As Heidegger puts it in *Being and Time*, *Being historical* or ‘historicality’ or ‘historicity’ (*Geschichtlichkeit*) is an a priori condition of being human (Heidegger, 1962). Or, as Merleau-Ponty famously put it, human being is “an historical idea, not a natural species” (Merleau-Ponty, 2012). Human existence then has to be understood in terms of its overall temporal and historical dimensions. At the same time, Husserl’s reduction is an attempt to break with the natural flow of this historicity (Husserl, 1970, p. 151). There is, so to speak, even such a thing as the ‘historical way’ to the reduction, that Husserl speaks about in 1937 and which he regarded as more primary than the way outlined in *Ideas I* (Moran, 2013b). Furthermore, as Husserl explored, all cultures live within horizons of familiarity and strangeness. Every life-world has dimensions of “home-world” and “alien-world” (Steinbock, 1995); everyone lives within horizons which lay out structures of normality and abnormality, harmony and surprise, yet these structural dimensions of experience rarely are foregrounded in intentional explication.

## **2. Phenomenology as a Tradition Self-Conscious of Traditionality**

### 3. Phenomenology as Primarily Intuitive, Descriptive and Essence-Seeking (*Wesensschau*)

Phenomenology essentially *is* intentional description, i.e. it aims to describe every kind of object or ‘objectual’ situation or state of affairs, in terms of its correlation with an apprehending subjectivity. In this sense, phenomenology is a philosophy of *correlation*. Full explanation requires taking the subjective point-of-view into consideration. In his *Crisis of European Sciences* Husserl claimed his real philosophical breakthrough came in 1898 when he realized that there was a “universal a priori of correlation between experienced object and manners of givenness” (Husserl, 1970, p. 166 n.). Every object must be understood not solely as it is ‘in itself’ but in necessary relation to the subjective acts that disclose it. Anything that is – whatever its meaning and to whatever region it belongs – is “an index of a subjective system of correlations”, Husserl, 1970, p. 165). Husserl writes:

The correlation between world (the world of which we always speak) and its subjective manners of givenness never evoked philosophical wonder (that is, prior to the first breakthrough of “transcendental phenomenology” in the *Logical Investigations*), in spite of the fact that it had made itself felt even in pre-Socratic philosophy and among the Sophists – though here only as a motive for skeptical argumentation. This correlation never aroused a philosophical interest of its own which could have made it the object of an appropriate scientific attitude. (Husserl, 1970, p. 165)

According to this a priori correlation, the manner in which entities in the world present themselves is always related to the subjective way of apprehending these entities. As Husserl puts it in his 1917 Inaugural Address to Freiburg University entitled “Pure Phenomenology: Its Method and Field of Investigation”:

To every object there correspond an ideally closed system of truths that are true of it and, on the other hand, an ideal system of possible cognitive processes by virtue of which the object and the truths about it would be given to any cognitive subject. (Moran & Mooney, 2002, p. 125)

Phenomenology then is a form of correlationism, although this is not to condemn it, as Quentin Meillassoux attempts to do in his recent attack on correlationism (Meillassoux, 2009). Correlationism aims to give a fuller account of ontology by including the subjective dimension, hence *phenomenological ontology*.

Phenomenology, broadly speaking, involves the careful, unprejudiced *description* of conscious, lived experiences (Husserl’s *Erlebnisse*), precisely according to the manner that they are experienced, without the imposition of external explanatory frameworks, whether these be drawn from the natural or social sciences, from religion, or even from common sense or ordinary language use. In *Being and Time*, Heidegger defines phenomenology as “to let that which shows itself be seen from itself in the very way in which it shows itself from itself” (Heidegger, 1962, p. 58). This formulation is close to that found in Husserl’s *Crisis of European Sciences*: “to take the conscious life, completely without prejudice, just as what it quite immediately gives itself, as itself, to be” (Husserl, 1970, p. 233). In this sense, phenomenology seeks to remain descriptively true to what conscious experience – understood in the widest possible sense – reveals itself to be to the disciplined observer. For Merleau-Ponty, the practice of phenomenological seeing is meant to disrupt the everyday. It aims at “disclosure of the world” (*révélation du monde*). In his *Phenomenology of Perception* (1945), he writes: “true philosophy entails learning to see the world anew” (Merleau-Ponty, 2012, p. lxxxv). The key to the methodology is the manner of this observation – which is not introspection (as it is often portrayed) but rather methodical reflective analysis that is seeking for invariants

and essential features structuring the flow of experience. Husserl often made somewhat grandiose claims about phenomenology as the “science of sciences”, and an ultimate presuppositionless science, and these claims can easily be challenged, but his overall vision of a foundational science that explored the entire life of consciousness uncovering its essential structures has an enduring appeal. This Husserl writes in his *Lectures on Passive Synthesis*:

But if one has learned to see phenomenologically and has learned to grasp the essence of intentional analysis, if one has—expressed in the form of the Goethian myth—found the way to the mother of knowledge, to its realm of pure consciousness in which all being arises constitutively and from which all knowledge as knowledge of beings has to fashion its ultimate comprehensible clarification, then one will initially make the quite astounding discovery that those types of lived experience are not a matter of arbitrary special features of an accidental life of consciousness, but rather that terms like “perception,” “memory,” “expectation,” etc., express universal, essential structures, that is, strictly necessary structures of every conceivable stream of consciousness, thus, so to speak, formal structures of a life of consciousness as such whose profound study and exact conceptual circumscription, whose systematic graduated levels of foundation and genetic development is the first great task of a transcendental phenomenology. It is precisely nothing less than the science of the essential shapes [*Wesensgestaltungen*] of consciousness as such, as the science of maternal origins. (Husserl, 2001a, pp. 365-6).

This is Husserlian phenomenology in a nutshell: “the science of the essential shapes [*Wesensgestaltungen*] of consciousness as such”.

Husserl is uncovering the formal structures of any consciousness whatsoever and thereby the various manners of givenness (*Gegebenheit* – what Heidegger called “the magic word” of phenomenology) through which the world is manifest to consciousness. Husserl is here self-consciously radicalizing Kant’s project of a critique of pure reason. For Husserl, if one can uncover the essential a priori structures of subjectivity (in its fullest sense – the whole of mental life or ‘life of spirit’ (*Geistesleben*), as he terms it), the source of all knowledge about the world, actual, possible, impossible, and so on, will be revealed. Husserl sought to uncover the essential nature of subjectivity itself, from its deepest layer of originary time-consciousness, up through the layers of embodied sensuous, perceptual and emotional life, to the very core of the self in its relations with other selves and to the horizoning, shared world. Husserl’s structural account of inner awareness of temporality (Husserl, 1990) is still a vital source for philosophical reflection.

Phenomenology has advocated both the ideal of a completely *presuppositionless* starting point for philosophical reflection and also the idea that all experiences are interpreted in terms of one’s linguistic, cultural and intentional horizons. Husserlian phenomenology, *ab initio*, claimed to be “presuppositionless” (*voraussetzunglos*: Husserl, 2001, vol. 1, p. 177). Thus he writes even in his Amsterdam Lectures of 1928:

We will refrain from any traditional prejudgments, even the most universally obvious ones of traditional logic, which already have perhaps taken from Nature unnoticed elements of meaning. We will hold ourselves resolutely to what phenomenological reflection presents to us as consciousness and object of consciousness, and purely to what comes to actual, evident self-givenness. In other words, we will interrogate exclusively the phenomenological experience, clearly and quite concretely thinking into a reflective experience of consciousness, without interest in determining concretely occurring facts. (Husserl, 1997)

#### **4. Presuppositionless Philosophy versus Hermeneutical Suspicion**

Husserl remained committed to presuppositionless as a methodological principle. On the other hand, Heidegger's development of hermeneutical philosophy, brought interpretation and 'pre-judgment' (*Vorurteil*) to the heart of the phenomenological discussion of understanding (*Verstehen*). Indeed, Husserl, too, in his late work, although he retained the goal of final 'clarification' (*Klärung*) did recognise the embeddedness of human being within their cultures and the manner in which their horizons were determined by their cultural embeddedness. As an example of how themes can return in different form (and without recognition of its *Vorwelt*), consider the phenomenon that, after several centuries of discussion of presupposition, prejudgement (*Vorurteil*), hermeneutics, conflict of interpretation, and so on, 'implicit bias' has now emerged in empirical psychology and has been, somewhat uncritically, embraced by contemporary analytic philosophy, without reference to the longer tradition. One wonders how Hans-Georg Gadamer's assertion that all understanding is misunderstanding, and Paul Ricoeur's hermeneutics of suspicion (Gadamer 1984) might sit with those enthusiastically detecting and claiming to neutralise if not eliminate implicit bias.

### 5. Intentionality as the Principal Theme of Phenomenology

Phenomenology begins from the recognition of intentionality: our subjective comportment and conscious behavior is essentially both meaning-apprehending and meaning-making, overall making sense of the world. Husserl speaks of *Sinngebung*, 'sense-giving', 'sense-bestowal'; 'sense-explication' (*Sinnauslegung*), captured in 'reflexion' (*Besinnung*). Furthermore, phenomenology maintains that intentionality is all pervasive: all aspects of live involve a coming together of subjective attitude and objective meaning. Husserl calls intentionality (*Intentionalität*) – elsewhere described as 'directedness' (*Gerichtetheit*) or 'aboutness' (*Meinung*) – the "principal theme" (*Hauptthema*) of phenomenology (Husserl, 2014, p. 161). Intentionality was reinserted into modern philosophy (recovered by the Scholastics) by Franz Brentano in his 1874 *Psychology from the Empirical Standpoint* (Brentano, 1995). In his effort to define the domain of psychological science, Brentano proposed that intentionality was the defining characteristic of all and only mental phenomena (Moran, 1996). Initially, Brentano conceived of intentionality rather narrowly as a relation between a mental act and its 'internal' object (dreaming about *winning the race* is an act of dreaming that is directed at a peculiar object *winning the race*, in this case a complex state-of-affairs). Husserl expanded this insight to make intentionality the meaning-endowing character of all lived experiences. Furthermore, Husserl treated intentionality not narrowly as a relation between a mental act and its object, and especially not as a kind of *representation* of the outer world in the inner mind (Drummond 2012; 1992; 1990), but as expressing the fundamental ways embodied subjects, interacting with other subjects, apprehend the world as senseful and respond actively generating further commitments of sense. Intentionality is essentially a claim about the *sensefulness* of experience based on the irreducible, essential interrelatedness between signifying consciousness and the surrounding world of significance. This interpretative activity need not be explicitly linguistic. In this regard, Husserl distinguishes between 'sense' (*Sinn*) (e.g. of non-linguistic perceptions) and 'meaning' (*Bedeutung*) that is expressed in language. Our embodied sensings, perceptions, feelings, moods, emotions, imaginings, thoughts, already have 'sense' prior to verbal articulation and may in fact be consciously experienced in an embodied manner. Our perceptual vista simply unfolds for us as we turn our eyes around to view the landscape; our hearing immediately experiences being in the midst of a noisy scene with car engines, slamming doors, human voices, barking dogs, or whatever. Even 'white noise' has a certain experiential significance – and is experienced as uncomfortable, distressing, and so on. All experiencing is experiencing of something and as something. We feel the day to be gloomy and we lack energy and enthusiasm. These are all forms of sense-disclosure. Furthermore, experiences occur against a backdrop of previously meaningful experiences,

in what Husserl calls the ‘harmoniousness’ (*Einstimmigkeit*) of our experience, which in turn motivates certain expectations of continuity, how the future situation will pan out, and so on. We have the sense of the world unfolding, going on, outside us and within us. Aron Gurwitsch especially developed Husserl’s account of the theme and its ‘thematic field’:

What is given in genuine sense experience presents itself as a member of a system, becoming what it phenomenally is by its relatedness to the system, by the role which it plays in it, by its significance for it. (Gurwitsch, 1966, p. 348)

Husserl’s and Gurwitsch’s accounts of perception had a major impact on Merleau-Ponty’s *Phenomenology of Perception* (1945; Merleau-Ponty, 2012), which is still extraordinarily influential in contemporary philosophy of mind.

Intentionality is widely discussed in analytic philosophy (Dennett, 1987; Searle, 1983) also, but phenomenology generally opposes immanent, representationalist, causal, and naturalist accounts of intentionality (Dreyfus 2000, 2007), and instead approach conscious acts and states as conferring ‘sense’ (*Sinn*), ‘meaning’, or ‘significance’ (*Bedeutung, Meinung*), in part constituted through embodied ‘comportment’ (*Verhalten*) and intersubjective ‘interaction’ (*Ineinandersein*), inhabiting a pre-given world already charged with significance (Dreyfus 1991; 2007a). Human beings cast a web of sense over their experiences, so even disturbed, pathological subjects imbue their fantasies and paranoid thoughts with senseful plausibility (Gendlin 1962). Indeed, for Husserl, so long as subjectivity exists, the idea of complete disorder, chaos, and senselessness (*Unsinn*), pure discordance, is literally unthinkable since, at the very least, the subjective flow of egoic consciousness will continue (Husserl 2014).

Husserl has a tiered or nested account of how intentional states are founded and interrelated. Perception is the primal or ‘originary’ form of intuiting, whereas memory and imagination are reproductive modifications of perception. Memory is a form of ‘calling to mind’ or ‘re-presenting’ (*Vergegenwärtigung*) that no longer has the distinctive bodily presence that characterizes perception. Imagining is yet another form of presenting which posits an object somewhat detached from perceptual surroundings. Furthermore, genuine representing or mental envisaging is, for phenomenology, just one peculiar intentional mode. Pictorial representing has its own structural character that is founded on the more direct perceptual intentionality. Empty or ‘signitive’ intendings, for Husserl, constitute the largest class of our conscious acts, and have a particular relevance in mathematics and scientific discourse where signs are manipulated in an empty way but ultimately, for Husserl must be grounded in fulfilled intuitions.

For Husserl, not just the modes of the intentional stance but, correlatively, the objects of intention are multiple. For instance, one can perceive objects and their properties, but also one can perceive actions, processes, relations, and complexes or events that Husserl calls “states of affairs” (*Sachverhalte*, Husserl 2001, vol. 2, 155). Perceptions found perceptual statements. Perceptual statements and judgments are, in turn, complex multi-layered acts – linguistic utterances select aspects of the overall perceived state of affairs (I can say that I see *the blackbird*, *the blackbird flying*, that *the bird is startled*, etc.) but never reproduce the intuitive fullness of the actual sensuous perception itself. In this regard, as Merleau-Ponty insists, perception is richer than thought; and conceptual judgements at best one-sidedly articulate what is implicit in our incredibly rich perceptions.

Furthermore, Husserl recognized that all objects that occupy space are apprehended in ‘profiles’ or ‘adumbrations’ (*Abschattungen*), and their essence is never exhausted by these profiles. An object is always further determinable and portends ever new contexts in which these refigured experiences can be fulfilled. Physical objects present with an ‘excess’



(*Überschuss*), whereby their determinate features are supplemented with a horizon of indeterminacy of features that can be explored in further perceptions (Hopp, 2011). In *Ideas I* (Husserl 2014), Husserl speaks of “modes of indeterminate suggestion and non-intuitive co-presence” (Husserl 2014, 183) that are wrapped up in the experience. Husserl writes:

No matter how completely we may perceive a thing, it is never given in perception with the characteristics that qualify it and make it up as a sensible thing from all sides at once. ... Every aspect, every continuity of single adumbrations, regardless how far this continuity may extend, offers us only sides. And to our mind this is not just a mere statement of fact: It is inconceivable that external perception would exhaust the sensible-material content of its perceived object; it is inconceivable that a perceptual object could be given in the entirety of its sensibly intuited features, literally, from all sides at once in a self-contained perception. (Husserl, 2001a, pp. 39–40)

Perception, for Husserl, anchors us to the world. Perception is the primal or ‘originary’ form of intuiting, accompanied by the firm conviction of things (Husserl’s *Urglaube*) being present thus and so. Perceptual objects are given to apprehension with full ‘bodily presence’ (Husserl: *leibhaftig*; Merleau-Ponty: *en chair et os*). Husserl writes:

[Perception] is what originally makes us conscious of the realities existing for us and “the” world as actually existing. To cancel out all such perception, actual and possible, means, for our total life of consciousness, to cancel out the world as objective sense and as reality accepted by us; it means to remove from all thought about the world (in every signification of this word) the original basis of sense and legitimacy. (Husserl, 1974, p. 26)

The perceived object is directly presented as a whole, even if it is really presented as possessing an indefinite number of further profiles that can be brought into view. Perception is also truth-giving, truth-grasping. It has a ‘filled’ character that provides the intuitive basis for our sense of truth. Husserl conceives of the structure of intentionality as essentially a dynamic movement of empty intentions towards fulfillment. Actual perception is the paradigm case of *fulfilled* intuition (albeit that the object always presents in profiles or shadings, *Abschattungen*). There are different degrees of fulfillment. Husserl also connects intentionality with the manner that objectivities have ‘self-giveness’ (*Selbstgegebenheit*) in experience, a phenomenon he often calls ‘self-evidence’ (*Evidenz*), and which he strongly distinguished from psychological feelings of certainty. Husserl sees intentionality and evidence as essentially correlated:

The concept of any intentionality whatever - any life-process of consciousness-of something or other - and the concept of evidence, the intentionality that is the giving of something-itself [*Selbstgebung*] are essentially correlative. (Husserl, 1969, p. 160)

Perfect evidence, for Husserl, involves the grasping of the object in full givenness and with the clarity and distinctness appropriate to it. For Husserl, intentionality is related to the self-giveness of experiences and self-giveness has the remarkable character of manifestness or what Husserl calls ‘phenomenality’. Thus he writes:

... one can say of every mental process that in it something is appearing to the particular I insofar as the I is somehow conscious of it. Accordingly, phenomenality [*die Phänomenalität*], as a characteristic that specifically belongs to appearing and to the

thing that appears, would, if understood in this broadened sense of the term, be the fundamental characteristic of the mental. (Husserl, 1997)

For Husserl, a physical object, apprehended phenomenologically, supports a potentially indefinite number of possible modes of access to it. Hence Husserl speaks of the intended object as “an idea in the Kantian sense” (Husserl, 2014, p. 284). Intentional objects already contain the possible modes of approach to them as a series of lawfully related noemata (Drummond & Embree, 1992). The aspectual shapes or modes of approach to the object can be visualized as ‘windows’ or avenues of approach to the object, set up in an essentially pre-determined way. Thus, in the *Amsterdam Lectures*, Husserl describes the noema of a house in a house-perception as opening onto an infinite horizon of other possible profiles of the house:

The question immediately arises as to how come it is evident that this pointing-ahead belongs to the phenomenon-in-consciousness? How come this horizon-consciousness refers us in fact to further actually unexperienced traits of the same <phenomenon>? Certainly, this is already an interpretation which goes beyond the moment of experiencing, which we have called the “horizon-consciousness,” which is, indeed, as is easily determined, completely non-intuitive and thus in and of itself empty. (Husserl, 1997, pp. 226-227)

In his mature works, Husserl expands his earlier analyses of the specific features of intentional objects to consider the non-objectual intentionality of horizons, fringes, that manifest themselves in the temporal flow of a unified “nexus of consciousness” (*Bewusstseinszusammenhang*, Husserl, 1969, p. 159) and ultimately form the life-world (*Lebenswelt*) that is ‘always already there’ but can never be objectified. In *Ideas I* Husserl defines the ‘horizon’ as “‘what is co-given’ but not genuinely” (Husserl, 2014, p. 77) and, in *Experience and Judgment*, he speaks of “the horizon of typical pre-acquaintance in which every object is pre-given” (Husserl, 1973, p. 150). Husserl believed he has made a genuine breakthrough with his concept of horizontal-intentionality (*Horizont-Intentionalität*), originally inspired by William James. Husserl’s ‘horizon-intentionality’ is later taken up by Gurwitsch and Merleau-Ponty to explicate the complex manners experiences are framed by temporal and other horizons that have their own kind of significance.

Intentionality covers the whole of conscious life; everything is an achievement or accomplishment of intentional consciousness: “... nothing exists for me otherwise than by virtue of the *actual and potential performance of my own consciousness* [*Bewusstseinsleistung*]” (Husserl, 1969, p. 234). Life is *intentional*, ‘*accomplishing*’ life with its potentially infinite horizons of intentional implication, uniting together into the collective experience known as *spirit*. As Husserl writes in the *Crisis*:

... conscious life is through and through an intentionally accomplishing life [*intentional leistendes Leben*] through which the life-world ... in part attains anew and in part has already attained its meaning and validity. All real mundane objectivity is constituted accomplishment in this sense, including that of men and animals (Husserl, 1970, p. 204).

The Husserlian account of lived experiences is extraordinarily rich. For phenomenology, not just perceptions and judgments (the traditional focus of epistemology) but also feelings, emotions (Vendrell Ferran 2008, 2015), acts of willing, moods, have a peculiar manifestness of their own which contributes to the manner in which consciousness engages with the world. For Husserl, intentional experiences form a seamless unity in the flow of consciousness but

## 6. The Intentionality of Emotions and Moods

nevertheless, there is a structural order of experiences.

Husserl's account of intentionality was further radicalized by Heidegger and by Merleau-Ponty's analyses of embodied, habitual and practical intentionality, focusing especially on pre-theoretical, habitual, practical behavior in an environment of already-given significance. Phenomenologists have seen the significance of moods as being ways in which the world is disclosed and also the manner in which one is disclosed to oneself. Thus Stephan Strasser claims that moods are "the primordial phenomenological characteristic of self-experiencing life" (Strasser, 1997, p. 121). Heidegger, in *Being and Time*, made significant contributions with his analyses of 'fundamental moods' (*Grundstimmungen*), disposition or 'state-of-mind' (*Befindlichkeit*), and the experience of 'thrownness' (*Geworfenheit*), finding oneself always thrown into a situation. Heidegger (1962) discussed fundamental moods such as anxiety (*Angst*) which he distinguished from fear, a relationship to something in particular, but has as its object one's whole being-in-the-world. One's whole being is experienced as groundless and unsupported. Similarly Aurel Kolnai (2004) developed an analysis of disgust, and Sartre develops this with his account of vertigo or nausea – where one experiences one's own absolute freedom. Vertigo is not so much the fear of falling as the fear of jumping. One's inner freedom is limitless and, in this sense, truly terrifying, opening up before one like a chasm. Sartre's phenomenological accounts of the experience of freedom have enduring value (Sartre 1979).

### **7. The First-Person Point of View and the 'Cartesian' Starting Point**

Part of the enduring appeal of phenomenology is that it respects the centrality and ineliminability of the *first-person* point of view, the first and primary standpoint from which humans encounter the world. Phenomenology argues that the objective view --what Merleau-Ponty calls, in his 1945 *Phenomenology of Perception*, "the view from nowhere" (*la vue de nulle part*, Merleau-Ponty, 2012)—is achieved only by abstracting from the original first-person stance with which humans engage the world.

All knowing requires the engagement of subjectivity and is the achievement of a first-person. This is why Husserl stresses his so called 'Cartesian' way of practicing phenomenology, which Husserl articulated especially in the *Cartesian Meditations*, lectures given in Paris in 1929 (Husserl, 1960). Subsequently, in the *Crisis*, Husserl identified at least three 'ways' of performing the reduction: the "Cartesian way" (Husserl, 1970, p. 154), the "way from psychology" (*Weg von der Psychologie*, Husserl, 1970, p. 208), and the way from the pregiven life-world (Husserl, 1970, p. 136). Husserl admires Descartes' attempt to overthrow the existing sciences and set forth the indubitable grounds for a first philosophy that will be "a science grounded on an absolute foundation" (Husserl, 1960, p. 1). Husserl also agrees with Descartes that the practice of philosophy must be a first-personal affair, and that the meditator must accept only his or her own existence as having an absolute indubitability. For Husserl, Descartes has uncovered the fundamental fact that all being is relative to consciousness and that the transcendental ego has absolute being (Husserl, 1967, p. 69). There is an undeniable priority to my self-experience. Phenomenology attends to whatever has immediate 'self-giveness' in the realm of self-conscious intuitive experience. One can turn the focus of one's attention away from matters in the world to what is given in the stream of one's own first-personal conscious experiencing. One can attend to the various manners of givenness of whatever is encountered within that stream of experiencing.

This requires the meditating self or ego to come to understand the essential nature of the self in general, and thereby modifying its own nature to include the possibility of other different natures (other egos). Husserl thought that this insight into the essence of egohood or selfhood belonged exclusively to conscious experiences of the form of the cogito and offered a unique starting point for gaining what he considered to be apodictic knowledge. Note, however, that

Husserl has no interest in what is uniquely subjective and idiosyncratic to my experiencing – he is interested in the essential form of self-experiencing as such. Of course, Husserl’s embrace of Cartesianism is now seen as just one strand in his complex layered set of approaches. Husserl in later years sought to embed this phenomenology of self-experience into a wider phenomenology of the experience of alterity.

As Husserl’s phenomenology developed, he increasingly emphasised that we are *incarnate*, that is to say, embodied, *situated*, *finite* human beings, already in the world and for whom the world has a given, taken-for-granted, ‘natural’ status. Husserl himself speaks of the ‘phenomenology of embodiment’ (*Phänomenologie der Leiblichkeit*). Similarly, he attended to the en-worlded nature of our experience. Much of Husserl’s puzzlement came from trying to figure out how it is that the lived world comes to have this *taken-for-grantedness*, *self-givenness* and *obviousness* in the natural attitude.

Phenomenology is interested not just in the first-person, but in the second-person, in the *alter ego*, both singular and plural, and the constitution of various forms of ‘I-Thou’ relations (Husserl speaks of the ‘*Ich-Du-Beziehung*’) and ‘we-intentionality’ (*Wir-Intentionalität*) in what is now known as social or collective intentionality. Indeed, some of the most fruitful discussions concerning the formation of the ‘we’ and ‘we-world’ (*Wir-Welt*), as Husserl calls it, have emerged in recent social philosophy in dialogue with phenomenology (Moran & Szanto, 2015). Phenomenology is, as Husserl put it, a science of subjectivity, but it also a science of subjectivity embodied and embedded and involved in the intersubjective constitution of objectivity and of the world.

Phenomenologists can begin from their own experience, but they are open to understanding other people’s experiences, as they are experienced (examples can be found in real life, in literature, or, simply, through the process of imaginative variation, a procedure Husserl himself practiced). In other words, close attention must be paid to the subject’s own account of their experiences, emotions, and how they form their general sense of matters that affect them. Phenomenology is replete with descriptive analyses of affective life, feelings, emotions, moods, existential concerns, and various dimensions of lived embodiment and more generally ‘being-in-the-world’ (*In-der-Welt-Sein*, Heidegger, 1962; *être-au-monde*, Merleau-Ponty, 2012).

One of phenomenology’s greatest and most useful discoveries is that of the ‘natural attitude’ (*die natürliche Einstellung*)—a term that Husserl uses from as early as 1906-7, but which emerges in print in *Ideas I* (Husserl, 2014, p. 48). In order to focus on the *manner of experiencing*, it necessarily involves interrupting natural, straight-forward experiencing. According to Husserl, in our everyday practices and routines, we are in a certain attitude (he calls it ‘the natural attitude’) towards things and towards the world, and somehow this is a state of self-forgetfulness. The world presents itself as simply there, given, available to us. Husserl’s insight (inspired in part by Avenarius’ discussion of non-dualistic experiential life in *Der menschliche Weltbegriff*) is that the ordinary, natural world that surrounds us on all sides, in which we live and move and have our being, is actually itself the *correlate* of a very powerful yet also quite specific and particular attitude: the *natural attitude*. In his 1935 *Vienna Lecture* Husserl defines an attitude as “a habitually fixed style of willing life (*Stil des Willenlebens*) comprising directions of the will or interests that are prescribed by this style, comprising the ultimate ends, the cultural accomplishments whose total style is thereby determined” (Husserl, 1970, p. 280). In his Kant Society lecture in 1974, he describes the natural attitude:

The natural attitude is the form in which the total life of humanity is realized in running its natural, practical course. It was the only form from millennium to

## 8. Intersubjectivity, Sociality, Worldhood

## 9. The Natural Attitude and its Suspension

millennium, until out of science and philosophy there developed unique motivations for a revolution. (Husserl, 1974, 20)

Disrupting the natural attitude and undermining its hold on us will be central to Husserl's practice of the phenomenological method. It is important also to emphasize that the scientific worldview which is now entwined with the natural worldview does not necessarily provide us with a grasp of reality and truth in the fullest sense.

### 10. The Phenomenological Reduction

In order to lay bare the intentional constitution of the natural attitude, phenomenologists apply the procedure of 'suspension', 'bracketing', or phenomenological reduction, to strip away the presuppositions embedded in the natural attitude. Husserl's procedure of *epoché* (a term he introduced in *Ideas I* to stand for the set of operations that excludes the naive, natural, objective attitude to the world) is an explicit radicalisation and modification of Cartesian doubt. Husserl interpreted Descartes as attempting a universal world-negation, whereas he himself sought not negation but rather *neutralisation* of worldly commitments. Breaking with the 'natural attitude' to prepare the mind for the scientific attitude is also part of Descartes' programme. Husserl says that his phenomenological *epoché* differs from Cartesian doubt (Husserl, 2014), in that he interprets the actual, historical Cartesian doubt as a kind of dogmatic scepticism, involving the dogmatic denial of the existence of the world, rather than as the more Pyrrhonian scepticism, which keeps existential judgements in suspense and remains uncommitted. Husserl wants to uncover the "genuine sense of the necessary regress to the ego" (Husserl, 1960, p. 6). His method is a treating of assumptions as mere "acceptance-phenomena", "inhibiting", "putting out of play" (Husserl, 1960, p. 20). In *Ideas I*, for instance, Husserl speaks of *epoché* as rigorous 'exclusion' (*Ausschaltung*) and an 'abstention' (*Enthaltung*) from the methods and propositions of the philosophical tradition (Husserl, 2014, p. 52). The *epoché* and the associated 'phenomenological reductions' are meant to allow the nature of the experiences to become manifest without distortion and to lead back (Latin: *re-ducere*) to the apprehension of the essential ('eidetic') structures, contents, and objects and how they interconnect with other forms of consciousness.

Husserl borrowed the *epoché* from the Greek sceptics – and it is noteworthy in the quotation above from the *Crisis* that Husserl speaks of the correlation between world and subjectivity as already arising under the mode of skepticism about knowledge already in the ancient Greeks. Indeed, Husserl thinks ancient Greek philosophers – a "few Greek eccentrics" (*ein Paar griechischen Sonderlingen*, Husserl, 1970, p. 289), as he puts it in the Vienna Lecture – were responsible for an entirely new attitude that broke with the prevailing mytho-poetic way of seeing the cosmos and led to the 'break into' (*Einbruch*) of the theoretical attitude which is the source of Western science (the only truly unbounded or infinite science). Heidegger, however, does not make explicit use of Husserl's phenomenological reduction, whereas Husserl thought it as essential to the point of claiming that anyone who misunderstood the importance of the reduction could not be doing phenomenology. Merleau-Ponty accepted a modified version of the phenomenological reduction. He famously said that the lesson of Husserl's reduction was that it could not be carried out to completion because it ran up against the indissoluble life-world – this "rupture can teach us nothing but the unmovitated springing forth of the world" (Merleau-Ponty, 2012, p. lxxvii).

Husserl's 'bracketing' is a complex procedure, demanding perpetual vigilance. In one sense, what it brings to light is the extraordinary web of assumptions and practical beliefs that make up what he calls 'the natural attitude'. It has not been universally adopted within phenomenology, nevertheless, something like this bracketing and suspension is necessary to allow phenomena to come into view and be apprehended in the least distorted manner.

Husserl is notorious for his conception that all “sense and being” (*Sinn und Sein*, Hua I 117; *Seinssinn*, I 118) is constituted by the transcendental ego. Nevertheless, the philosophical idea of the ‘transcendental ego’ can be rehabilitated and given a reasonably plausible reconstruction. For Husserl, the ego is a dynamic entity, a self that grows and develops, sediments properties that become convictions, and has a history. When a decision is made, this is not just a passing episode but becomes a permanent possession (even if it is later renounced). Husserl characterises the ‘pure’ or ‘transcendental’ ego as an “Ego-pole” (*Ichpol*, Husserl, 1960, p. 68) or centre of experiences, including all actions and affections, as the “substrate of habitualities” (Husserl, 1960, p. 69). The ego is not an empty ‘pole’ that sends and receives conscious ‘rays’, it is a concrete living self, an individual, an “abiding personal ego” (*als stehendes und bleibendes personales Ich*, Husserl, 1960, pp. 66-67), a monad. It collects its experiences together into an enduring unity through the performance of various complex syntheses and becomes a *self*, with convictions, values, an outlook, a history: “The ego constitutes himself for himself in, so to speak, the unity of a history” (Husserl, 1960, p. 75). The transcendental ego, moreover, necessarily functions within the open field of all possible egos – transcendental intersubjectivity (akin to the open field of language which is an intersubjective network that is activated in each speaker). Husserl writes:

But each soul also stands in community [*Vergemeinschaftung*] with others which are intentionally interrelated, that is, in a purely intentional, internally and essentially closed nexus [*Zusammenhang*], that of intersubjectivity. (Husserl, 1970, p. 238)

Merleau-Ponty correctly emphasized this point – transcendental subjectivity is already an intersubjectivity, but that is not to say that this statement resolves the extremely difficult problems concerning the constitution of the transcendental ego. Indeed, many phenomenologists, beginning with Heidegger, abandoned the essentially Kantian language of transcendental subjectivity in favor of discussion of the conditions of the possibility of historical human existence. However, I believe much can still be learned from the mature Husserl’s meditations on transcendental subjectivity especially in relation to the constitution of its enviroing world.

Classical phenomenology has many detailed analyses of the structures of conscious life and of the ‘life-world’ – the ordinary, everyday, pre-scientific world that we all inhabit and which provides a unified backdrop for all our intentional activities across history. The mature Husserl thus enlarged the scope of phenomenology by suggesting new approaches that moved beyond the exploration of subjectivity, specifically, in the *Crisis*, “by *inquiring back from the life-world*”, which is, of course, also the accepted, social and communal world that is, in Heidegger’s words, “always already there” (*immer schon da*). It is precisely the experience of intersubjectivity and the concept of a horizontal world as backdrop to all experience that requires a new kind of phenomenology of collective cultural experience. Heidegger’s distinction between inauthentic and authentic selfhood, and especially his discussion of ‘the one’, *das Man*, anyone, also contributes to the understanding of a kind of participatory stance that is anonymous and not personally defined in which we stand in the world. The life-world is an extra-ordinary and enduring innovation. Hans-Georg Gadamer has written:

In Husserl’s later work the magic word *Lebenswelt* (lifeworld) appears—one of those rare and wonderful artificial words (it does not appear before Husserl) that have found their way into the general linguistic consciousness, thus attesting to the fact that they bring an unrecognized or forgotten truth to language. So the word “*Lebenswelt*”

## 11. Transcendental Subjectivity

## 12. The Relativity and Universality of the Life-World

has reminded us of all the presuppositions that underlie all scientific knowledge. (Gadamer, 1998, p. 55)

Many twentieth-century European philosophers—including Jan Patočka (2016), Jürgen Habermas (1987), and Hans Blumenberg (1986) – have taken over Husserl’s conception of the life-world. The life-world is a world of cumulative tradition acquired through what Husserl calls *sedimentation* (*Sedimentierung*, Husserl, 1970, p. 362), according to which certain earlier experiences become passively enfolded in our on-going experience, just as language retains earlier meanings in its etymologies. As Husserl says in the “Origin of Geometry”, “cultural structures, appear on the scene in the form of tradition; they claim, so to speak, to be ‘sedimentations’ (*Sedimentierungen*) of a truth-meaning that can be made originally self-evident” (Husserl, 1970, p. 367). Indeed, Husserl speaks of ‘sedimentation’ as “traditionalization” (Husserl, 1970, p. 52). There never has been such an investigation of the ‘life-world’ as “subsoil” (*Untergrund*) for all forms of theoretical truth (Husserl, 1970, p. 124). This science of the life-world would be descriptive of the life-world in its own terms, bracketing conceptions intruding from the natural and cultural sciences. In this sense, Husserl speaks of an “ontology of the life-world” (*Ontologie der Lebenswelt*, see *Crisis* § 51). In the Kant Society lecture Husserl speaks of the results of the phenomenological method as follows:

The world took on an infinite wideness as soon as the actual life-world, the world in the “how” of the givenness of mental process [*die wirkliche Lebenswelt, die Welt im Wie der Erlebnisgegebenheit*], was observed. It took on the whole range of the manifold subjective appearances, modes of consciousness, modes of possible position-taking; for it was, for the subject, never given otherwise than in this subjective milieu, and in purely intuitive description of the subjectively given there was no in-itself that is not given in subjective modes of the for-me or for-us, and the in-itself itself appears as a characteristic in this context and has to undergo therein its clarification of sense. (Husserl, 1974, p. 11).

Husserl contrasts the practically experienced life-world with the world as given in objective science. He writes in *Ideas II*:

In ordinary life [*im gewöhnlichen Leben*] we have nothing whatever to do with nature-objects [*Naturobjekten*]. What we take as things are pictures, statues, gardens, houses, tables, clothes, tools, etc. They are value-objects [*Wertobjekte*] of various kinds, use-objects [*Gebrauchsobjekte*], practical objects. They are not objects which can be found in natural science. [*Es sind keine naturwissenschaftlichen Objekte*] (Husserl, 1989, p. 29)

Husserl in fact thinks the concept of an objective ‘world in itself’ that is considered independent of all subjectivity is a peculiar theoretical construction – useful in the sciences but one-sided in terms of metaphysics:

The contrast between the subjectivity of the life-world and the “objective,” the “true” world, lies in the fact that the latter is a theoretical-logical substraction [*Substruktion*], the substraction of something that is in principle not perceivable, in principle not experienceable in its own proper being, whereas the subjective, in the life-world [*das lebensweltlich Subjektive*], is distinguished in all respects precisely by its being actually experienceable. (Husserl, 1970, p. 127)

Phenomenology as a descriptive practice opposes any form of analysis that reduces the unified nexus of conscious flowing life to a collection of isolated and atomistic mental experiences. Lived experiences are the starting point – but these cannot be narrowly thought of as, for example, stimuli (Quine’s ‘noises and scratches’) or as sense data. Experience is a multi-modal sensory intertwining and interweaving, not just of the different senses (touch, sight, hearing, smell, taste) and our bodily proprioception, but also as experience filtered by mood and inflected by emotion and various cognitive attitudes. Phenomenology has always recognized that perception is a multimodal, embodied *achievement*, an intertwining of various sense modalities (sight, touch) and of bodily proprioceptive movements (called ‘kinaestheses’ in the psychological terminology of the time), such as eye movements, hand movements, movements of the body (to get nearer, get a better grip, look more closely), involving looking over the object, pointing, grasping, moving around the object, and so on. The body is in the world as the “heart is in the organism”, Merleau-Ponty declares (Merleau-Ponty, 2012, p. 209). Embodied perception brings the visible, tangible, spatial and material world alive. There is a very fine attunement between lived body and world. Moreover, disorders of the embodied relation to the world involve a significant modification of the way the world appears, as in pathological cases such as the case of the brain damaged former soldier Schneider, discussed by Merleau-Ponty (2012, pp. 105-140). More recently, for example, Louis A. Sass and Elizabeth Pienkos (2013), who have studied delusion from the phenomenological perspective, have identified necessary structural features of delusions, e.g., mutations of time, space, causality, self-experience, or sense of reality, that challenge accounts of delusion as a belief mode or doxastic state. They discuss the delusional mood which involves feelings of strangeness and tension, and a sense of tantalizing yet ineffable meaning. This simply illustrates how phenomenology is being used in descriptions of embodied experience in a wide range of settings.

Phenomenology, moreover, recognises that persons are in part *constituted through their feelings, emotions and moods*. Moods ‘colour’ emotions (Husserl, Stein) can intensify, weaken, extend, curtail emotions. One’s world is filtered through emotions and moods (Scheler, Heidegger). Emotions are not just felt by and expressed in the body (e.g. facial expressions, smiling) but *the whole body* is inhabited emotionally: e.g. the whole body can be tense; a way of walking can be nervous, one’s social world can be experienced as anxiety-filled, and so on.

Phenomenology is also very useful for exploring the notions of foreground, background and horizon in experiences. The philosophical discussion of the intentional greatly benefits from being broadened to the discussion of overall being-in-the-world. There is always in any experience a combination of present and absent moments, to see the front side is at the same time to have an empty intending apprehension of the rear side of an object in the form of a determinable indeterminacy. Furthermore, as Heidegger emphasizes, what one sees is constrained by one’s background assumptions and by the overall context of the ‘environment’ (*Umwelt*) and life-world. A craftsperson can identify the appropriate tools for a particular task and immediately sees what needs to be done. The gardener identifies some plants as *flowers* and others as *weeds*, not based on botanical classification but based on their usefulness and appropriateness in the context of gardening (itself a cultural and historically inflected practice). This is what Heidegger means when he says we first encounter objects in a practical way as ‘ready-to-hand’ (*zuhanden*) rather than merely as ‘present-at-hand’ (*vorhanden*), objects merely theoretically apprehended, shorn of their context of usefulness (Heidegger, 1962; Dreyfus, 1991). For Heidegger, humans are already embedded (*‘immer schon da’*) in a pre-formed, and largely intuitively-apprehended, historical and cultural world. Moreover, they are oriented primarily to the future, and the past is always implicitly interpreted in terms of this future project. Human beings live in an *Umwelt* with which one interacts, very much like one “lays down a path in walking” as Francisco Varela put it. Or like the manner in which a spider

### 13. Some Enduring Features of the Phenomenological Approach



generates its web from the silk thread produced within its own body from the spinnerets in its abdomen. Against Kant, who denied the experience of the totality of the world as such, Husserl defends the consciousness of world as a genuine phenomenon:

I am continually conscious of individual things in the world, as things that interest me, move me, disturb me, etc., but in doing this I always have consciousness of the world itself, as that in which I myself am, although it is not there as is a thing, does not affect me as things do, is not, in a sense similar to things, an object of my dealings. If I were not conscious of the world as world, without its being capable of becoming objective in the manner of an [individual] object, how could I survey the world reflectively and put knowledge of the world into play, thus lifting myself above the simple, straightforwardly directed life that always has to do with things? (Husserl, 1970, p. 251)

One of the challenges of phenomenology is how to express its descriptive findings. The ability to make fine discriminations must be matched with an equal ability to translate these discriminations into appropriately fine-grained linguistic communication (for a critique of phenomenology as lacking analytic rigor, see Searle, 2005). Husserl himself recognized this problem but did not address it centrally at least until some of his later writings, for instance his essay “On the Origin of Geometry” (Husserl, 1970), where he accorded to written language an enormously important role in fixing the meanings of ideal objectivities such as occur in mathematics so that they can be accessed as the same over and over again. For Heidegger, however, the issue of language became inescapable and marked a major turning in his conception of phenomenology and its possibilities. Subsequent phenomenology (Jacques Derrida, for instance, in so far as his work is motivated by phenomenology and continues to work within the phenomenological *epoché*, as he himself has attested) has had to grapple with the complexity of the relationship between language and experience in ways that have frequently challenged many of Husserl’s assumptions.

Phenomenology is an enduring resource for philosophers interested in the first-person perspective, the subjective experience of embodied living, perceiving, and knowing, the affective life with its disclosive moods, the existential experience of temporality and the complex co-intentionalities involved in sociality, empathy and intersubjective communal life. Phenomenology still has much to contribute in the understanding of life-world, the experience of the other, and the complex phenomena of interculturality.

### REFERENCES

- Blumenberg, H. (1986). *Lebenszeit und Weltzeit*. Frankfurt am Main: Suhrkamp;
- Brentano, F. (1995). *Psychology from an Empirical Standpoint*. [1874]. Trans. A.C. Rancurello, D.B. Terrell, and L.L. McAlister. London: Routledge;
- Brentano, F. (1995a). *Descriptive Psychology*. Ed. by R.M. Chisholm and W. Baumgartner. Trans. B. Müller. London: Routledge;
- Dahlstrom, D. O., Elpidorou, A., & Hopp, W., Eds. (2015). *Philosophy of Mind and Phenomenology: Conceptual and Empirical Approaches*. New York & London: Routledge;
- Dennett, D. (1987). *The Intentional Stance*. Cambridge, MA: MIT Press;
- Durt, Ch., Fuchs, Th., & Tewes, Ch., Eds. (2017). *Embodiment, Enaction, and Culture. Investigating the Constitution of the Shared World*. Cambridge, MA: MIT Press;
- Dreyfus, H. L. (2007a). The Return of the Myth of the Mental. *Inquiry*, 50, 4, 352-365;
- Dreyfus, H. L. (2007). Intelligence without Representation Merleau-Ponty’s Critique of Mental Representation: The Relevance of Phenomenology to Scientific Explanation, *Phenomenology and the Cognitive Sciences*, 1, 367-383;

- Dreyfus, H. L. (2000). A Merleau-Pontyan Critique of Husserl's and Searle's Representationalist Accounts of Action. *Proceedings of the Aristotelian Society*, 100 (1), 287-302;
- Dreyfus, H. L. (1991). *Being-in-the-World. A Commentary on Heidegger's Being and Time, Division One*. Cambridge: MIT Press;
- Drummond, J.J. (2012). Intentionality without Representationalism. In D. Zahavi (Ed.). *The Oxford Handbook of Contemporary Phenomenology*. Oxford: Oxford University Press;
- Drummond, J.J. (1990). *Husserlian Intentionality and Non-foundational Realism: Noema and Object*. Dordrecht: Springer;
- Drummond J. J., & Embree, L. E., Eds. (1992). *The Phenomenology of the Noema*. Dordrecht: Springer;
- Gadamer, H.-G. (1998). The Ideal of Practical Philosophy. In H.- G. Gadamer, *Praise of Theory: Speeches and Essays* (pp. 50-61). Trans. Chris Dawson. New Haven: Yale University Press;
- Gadamer, H.-G. (1984). The Hermeneutics of Suspicion. *Man and World*. 17, 3-4, 313-323;
- Gallagher, S., & Zahavi, D. (2012). *The Phenomenological Mind*. 2<sup>nd</sup> Edition. London: Routledge;
- Gendlin, E. T. (1962). *Experiencing and the Creation of Meaning: A Philosophical and Psychological Approach to the Subjective*. Evanston: Northwestern University Press;
- Giorgi, A. (2009). *The Descriptive Phenomenological Method in psychology: A modified Husserlian approach*. Pittsburgh: Duquesne University Press;
- Giorgi, A. (1997). The Theory, Practice, and Evaluation of the Phenomenological Method as a Qualitative Research Procedure. *Journal of Phenomenological Psychology*, 28 (2), 235-260;
- Gurwitsch, A. (1996). *Studies in Phenomenology and Psychology*. Evanston: Northwestern University Press;
- Habermas, J. (1987). *The Theory of Communicative Action. Volume Two: Lifeworld and System: A Critique of Functionalist Reason*. Trans. Thomas McCarthy. Cambridge: Polity;
- Heidegger, M. (1962). *Being and Time*. Trans. John Macquarrie and Edward Robinson. Oxford: Basil Blackwell;
- Hopp, W. (2011). *Perception and Knowledge: A Phenomenological Account*. Cambridge & New York: Cambridge University Press;
- Husserl, E. (2014). *Ideas for a Pure Phenomenology and Phenomenological Philosophy. First Book. General Introduction to Pure Phenomenology*. Trans. Daniel O. Dahlstrom. Indianapolis: Hackett;
- Husserl, E. (2002). 'Philosophy as Rigorous Science', trans. Marcus Brainard, *The New Yearbook for Phenomenology and Phenomenological Philosophy* 2, 249-295;
- Husserl, E. (2001a). *Analyses Concerning Passive and Active Synthesis. Lectures on Transcendental Logic*. Trans. Anthony J. Steinbock. Husserl Collected Works Volume IX. Dordrecht: Kluwer;
- Husserl, E. (2001). *Logical Investigations*, 2 Vols. Trans. J.N. Findlay. Ed. with a New Introduction by Dermot Moran and New Preface by Michael Dummett. London & New York: Routledge;
- Husserl, E. (1997). *Psychological and Transcendental Phenomenology and the Confrontation with Heidegger (1927-31), The Encyclopaedia Britannica Article, The Amsterdam Lectures "Phenomenology and Anthropology" and Husserl's Marginal Note in Being and Time, and Kant on the Problem of Metaphysics*. Trans. T. Sheehan and R.E. Palmer. Collected Works VI. Dordrecht: Kluwer;
- Husserl, E. (1990). *On the Phenomenology of the Consciousness of Internal Time*. Trans. J.B. Brough. Collected Works IV. Dordrecht: Kluwer;
- Husserl, Edmund (1989). *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy. Second Book. Studies in the Phenomenology of Constitution*. Trans. R. Rojcewicz and A. Schuwer. Dordrecht: Kluwer Academic Publishers;
- Husserl, E. (1974). Kant and the Idea of Transcendental Philosophy. Trans. T. E. Klein and W. E. Pohl. *Southwestern Journal of Philosophy*, 5, 9-56;
- Husserl, E. (1973). *Experience and Judgment: Investigations in a Genealogy of Logic*. Revised and Edited by L. Landgrebe. Trans. J.S. Churchill and K. Ameriks. London: Routledge and Kegan Paul;

- Husserl, E. (1970). *The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy*. Trans. D. Carr. Evanston, IL: Northwestern University Press;
- Husserl, E. (1969). *Formal and Transcendental Logic*. Trans. D. Cairns. The Hague: Nijhoff;
- Husserl, E. (1960). *Cartesian Meditations: An Introduction to Phenomenology*. Trans. D. Cairns. The Hague: Martinus Nijhoff;
- Kelly, S. D. (2000). *The Relevance of Phenomenology to the Philosophy of Language and Mind*. Studies in Philosophy. London: Routledge;
- Kolnai, A. (2004). *On Disgust*. Ed. Carolyn Korsmeyer and Barry Smith. Chicago: Open Court;
- Lau, K.-Y. (2016). *Phenomenology and Intercultural Understanding. Toward a New Cultural Flesh*. Contributions to Phenomenology 87. Dordrecht: Springer;
- Meillassoux, Q. (2009). *After Finitude*. London: Continuum;
- Merleau-Ponty, M. (2012). *The Phenomenology of Perception*. Trans. Donald A. Landes. London: Routledge;
- Moran, D., & Szanto, Th. Eds. (2015). *Discovering the We: The Phenomenology of Sociality*. London & New York: Routledge;
- Moran, D. (2014). 'What Does Heidegger Mean by the Transcendence of Dasein?' *International Journal of Philosophical Studies*, 22 (4), 491-514;
- Moran, D. (2013). Intentionality: Some Lessons from the History of the Problem from Brentano to the Present. *International Journal of Philosophical Studies*, 21 (3), 317-358;
- Moran, D. (2013a). "Let's Look at It Objectively": Why Phenomenology Cannot be Naturalized. *Phenomenology and Naturalism, Philosophy Supplementary Volume 72*, 89-115;
- Moran, D. (2013b). From the Natural Attitude to the Life-World. In L. Embree & T. Nenon (Eds.), *Husserl's Ideen*. Contributions to Phenomenology 66 (pp. 105-124). Dordrecht: Springer, 2013;
- Moran, D. (2008). Husserl's Transcendental Philosophy and the Critique of Naturalism. *Continental Philosophy Review*, 41, 4, 401-425;
- Moran, D. (2000a). Heidegger's Critique of Husserl's and Brentano's Accounts of Intentionality'. *Inquiry* 43 (1), 39-65;
- Moran, D. (2000). *Introduction to Phenomenology*. London & New York: Routledge;
- Moran, D. (1996). Brentano's Thesis. *Proceedings of the Aristotelian Society*, 70, 1-27;
- Moran, D., & Mooney, T., Eds. (2002). *The Phenomenology Reader*. London & New York: Routledge;
- Mulligan, K., Ed. (1987). *Speech Act and Sachverhalt: Reinach and the Foundations of Realist Phenomenology*. Reidel;
- Noë, A. (2005). *Action in Perception*. Cambridge, MA: MIT Press;
- Patočka, J. (2016). *The Natural World as a Philosophical Problem*. Evanston: Northwestern University Press;
- Petitot, J., & Varela, F. J., et al., Eds., (1999). *Naturalizing Phenomenology*. Stanford: Stanford University Press;
- Pfänder, Alexander (1967). *Phenomenology of Willing and Other Phaenomenologica*. Trans. Herbert Spiegelberg. Evanston: Northwestern University Press;
- Salice, A. (2009). *Urteile und Sachverhalte. Ein Vergleich zwischen Alexius Meinong und Adolf Reinach*. München: Philosophia Verlag;
- Salice, A., & Schmid, H.-B., Eds. (2016). *The Phenomenological Approach to Social Reality*. Dordrecht: Springer;
- Sartre, J.-P. (1979). *Being and Nothingness*. Trans. Hazel Barnes. London: Routledge;
- Sass, L. A., & Pienkos, E. (2013), Delusion: The Phenomenological Approach. In K.W.M. Fulford, & M. Davies, et al. (Eds.). *The Oxford Handbook of Philosophy and Psychiatry*. Oxford: Oxford University Press;

- Schmid, H.-B. (2009). *Plural Action. Essays in Philosophy and Social Science*. Contributions to Phenomenology Vol. 58. Dordrecht: Springer;
- Searle, J. R. (1983). *Intentionality*. New York: Oxford University Press;
- Searle, J. R. (2005). The Phenomenological Illusion. In M. E. Reicher & J. C. Marek (Eds.). *Experience and Analysis. Erfahrung und Analyse*. Proceedings of the 27th International Wittgenstein Symposium, Kirchberg, Austria, August 8-14, 2004 (pp. 317-336). Vienna, 2005;
- Smith, J.A., Flowers, P., & Larkin, M. (2009). *Interpretative Phenomenological Analysis: Understanding Method and Application*. London: Sage;
- Steinbock, A. (1995). *Home and Beyond: Generative Phenomenology after Husserl*. Evanston: Northwestern University Press;
- Strasser, S. (1977). *Phenomenology of Feeling: An Essay on the Phenomena of the Heart*. Pittsburgh: Duquesne University Press;
- Straus, E. W. (1966). The Upright Posture. In *Phenomenological Psychology*. New York: Basic Books, 137-165;
- Szanto, Th. (2014). Social Phenomenology: Husserl, Intersubjectivity, and Collective Intentionality. *International Journal of Philosophical Studies*, 22 (2), 296-301;
- Szanto, Th. (2012). *Bewusstsein, Intentionalität und Mentale Repräsentation. Husserl und die Analytische Philosophie des Geistes*. Berlin-New York: De Gruyter;
- Thompson, E. (2007). *Mind in Life: Biology, Phenomenology, and the Sciences of Mind*. Cambridge, MA: Harvard University Press;
- Vendrell Ferran, I. (2015). The Emotions in Early Phenomenology. *Studia Phaenomenologica*, 15, 349-374;
- Vendrell Ferran, I. (2008). *Die Emotionen. Gefühle in der realistischen Phänomenologie*. Berlin: Akademie;
- Waldenfels, B. (2007). *The Question of the Other*. Hong Kong: The Chinese University Press / New York: SUNY Press;
- Zahavi, D. (2013). Naturalized Phenomenology: A desideratum or a category mistake?. In H. Carel & D. Meacham (Eds.). *Phenomenology and Naturalism*. Royal Institute of Philosophy Supplement, No. 72 (p. 23-42). Cambridge: Cambridge University Press;
- Zahavi, D. (2010). Naturalized Phenomenology. In S. Gallagher & D. Schmicking (Eds.). *Handbook of Phenomenology and Cognitive Science* (pp. 2-19). Dordrecht: Springer;
- Zahavi, D. (2004). Phenomenology and the Project of Naturalization. *Phenomenology and the Cognitive Sciences*, 3, 331-347.



---

SARA HEINÄMAA

Academy of Finland, University of Jyväskylä  
sara.heinamaa@jyu.fi

---

# TWO WAYS OF UNDERSTANDING PERSONS: A HUSSERLIAN DISTINCTION

---

## *abstract*

*This paper clarifies the distinction that Edmund Husserl makes between two different ways of understanding other persons, their actions and motivations: the experiential or empirical way, on the one hand, and the genuinely or authentically intuitive way, on the other hand. The paper argues that Husserl's discussion of self-understanding clarifies his concept of the intuitive understanding of others and allows us to explicate what is involved in it: not just the grasping of the other's actual motivations of action but also the grasping of her motivational possibilities. The paper ends by discussing the dynamic character of the personal subject.*

---

## *keywords*

*understanding others, empirical, intuitive, possibilities, person*

---

Husserl's mature ethics is personalistic in the sense that it ties the categorical imperative to individual persons (Hua37, 162, 246; Hua27, 31, 45; cf. Melle 2007; Hart 2006; Trincia 2007; Donohoe 2010; Siles i Borràs 2011; Beyer 2012; Heinämaa 2014; Crespo 2015). More precisely, Husserl defines the imperative, not in relation to types of actions or acts, but in relation to whole persons, their personal capacities<sup>1</sup> of acting, valuing and thinking and their ways of being motivated. In his late lectures on ethics, he formulates the imperative as follows: "Do from now on and without hesitation always the best, *your best*, grasp it in norm-cognition and will it in norm-conscious volition" (Hua37, 253, emphasis added). He explains:

What I must is determined by the 'I can', and what I can differs from what any other can [...] My best is, more precisely, determined by my past and my present, and my future is not totally without preference. My whole life lies in front of me, and in front of me lies the surrounding world that extends itself around me. What I can accomplish there operates as the basis for my deliberation, and the best that I can accomplish there now and in the future is my obligation [*Gesolltes*], the obligation of an individual (Hua37, 252–253; cf. 162; Hua4, 268/280).<sup>2</sup>

In the second volume of *Ideas*, Husserl gives murder and theft as examples of practical possibilities that are relevant in moral and ethical considerations, and he distinguishes two ways of considering such actions. Only one of these ways touches upon the person as an individual and belongs to ethics.

On the one hand, I can represent myself as committing an act of murder or a theft, but, on the other hand, such a representation is inauthentic if it lacks intuitive evidence concerning my

---

1 In paragraph §59 of the second volume of his *Ideas*, Husserl defines the person as a system of spiritual faculties and abilities (*Vermögen*) of the form "I can" (Hua4, 253–257/266–269). He develops this conception in his working notes and manuscripts on intersubjectivity published in the three volumes of his completed works, *Husserliana* XII–XV. In the following I will call these volumes "the Intersubjectivity Volumes" and distinguish between the first (Hua13), the second (Hua14) and the third (Hua15).

2 In the second volume of *Ideas*, we read: "A human being, however, has an individual kind [*Art*], and each human being has a different one. According to the universal [*dem allgemeinen nach*], she is a human being, but her kind as her character, her person, is a unity, constituted in her course of life as a subject of position-takings, a unity of multifarious motivations based upon multifarious presuppositions" (Hua4, 274/286–287, translation modified; cf. Hua 4, 270–271/283–284, 278–279/290–292, 297/311; Hua15, 118).

personality as a subject of actions, motivated by experiences and experienced objectivities. I can imaginatively represent myself as killing somebody or taking his or her possessions, but in so far as the motivations for such actions contradict my personality, my personal ways of being motivated and being able to become motivated, the reflection is abstract and proceeds in isolation from my motivational grounds and does not concern me as a person. Such possibilities are merely verbal or imaginary constructions and not my genuine possibilities of action. What is lacking is “the original [*ursprünglich*] consciousness of being able to do this action or of having the power for this action” (Hua4, 265/277).

In this context, Husserl points out that even in case of fictitious action self-consciousness is always “an originary [*originär*], non-neutralized consciousness” (Hua4, 265/277). More precisely, the consciousness of who I am, that is how I let myself be motivated, is not neutralized in the phantasizing self-inspection but, on the contrary, guides my imaginative variation of my behavior, my past deeds and my forthcoming actions and all other practical possibilities.

This means that our possibilities of conducting ethical lives or shaping our lives as ethical subjects depend on our capacities of understanding ourselves and others as persons, that is as subjects of motivated actions. More concretely, our possibilities of living ethical lives and judging ethically depend on our understanding of persons, their lived environments, their abilities and faculties and practical possibilities, their ways of acting and being motivated. We find these concepts – the concepts of person, motivation and understanding – developed by Husserl in the second volume of *Ideas* and related manuscripts. In this paper, I confine myself to the task of explicating Husserl’s concept of understanding persons as he uses it in these texts.

In the second volume of *Ideas* and in his working notes on intersubjectivity, Husserl distinguishes between two different ways of understanding persons, their actions and motivations: an experiential or empirical(-inductive) way, on the one hand, and the genuinely or authentically intuitive way, on the other hand.<sup>3</sup>

In paragraph §60 of the second volume of *Ideas*, Husserl asks: “What is happening when the character [*Charakterartung*] of a person suddenly lights up for us through some one or other of his glances, positions, or expressions; when we, so to say, ‘gaze into an abyss;’ when the ‘soul’ of the person suddenly ‘opens itself up;’ when we ‘fathom wondrous depths;’ etc.? What sort of ‘understanding’ is that?” (Hua4, 273/286).

The metaphors of depth and abyss are repeated in many other contexts that deal with the task of understanding ourselves and others.<sup>4</sup> In the later work *The Crisis*, Husserl cites Heraclitus’s

## 1. Empirical Understanding of Others

---

<sup>3</sup> In the first Intersubjectivity Volume (Hua13), Husserl uses the term “empathetic reflection”.

<sup>4</sup> Husserl uses the metaphors of surface and depth for two purposes, one constitutional-analytic and the other motivational-descriptive. In the analysis of the constitution of the ego as a person, these two senses combine. By the constitutional-analytical context, I mean Husserl’s discussion of the constitution of sense and his analyses of the dependency-relations and layers between different senses. This involves descriptions of levels of constitution as differing in depth and being in relations of one-sided and/or mutual dependency (e.g. Hua6, 96–97/94, 118/116, 121–123/118–120). By the motivational-descriptive context I mean Husserl’s discussion of motivations, that is the ways in which experiences are motivated by other experiences and what Husserl takes to be the kernel or core of personhood (the ego). Here the point is that we are motivated by our previous and present experiences and also by our perceptions and understanding of others. Husserl makes clear that whatever has motivational power over us must be experienced by us in one way or other: “What [...] does not stand over against me in my lived experiences, in my representing, thinking and acting, as the represented, perceived, remembered, thought etc., does not ‘determine’ me as a spirit. And what is not intentionally included in my lived experiences, even if unattended or implicit, does not motivate me, not even unconsciously” (Hua4, 231/243, cf. 185–186/195). Husserl does not reject the idea of unconsciousness but sets out to redefine the concept by horizontality, anonymity, association, striving and drives (e.g. Hua6, 240; Hua4, 222; Hua11,



famous statement: “You will never find the boundaries of the soul, even if you follow every road; so deep is its ground,” and he elaborates on this by characterizing the abyssal depth of the human soul as follows: “Indeed, every ‘ground’ that is reached points to further grounds, every horizon opened up awakens new horizons, and yet the endless whole, in its infinity of flowing movement, is oriented toward the unity of one meaning” (Hua6, 173/170).

So, in Husserl’s descriptions of the personal self and its motivational structures, the metaphors of wells and sources parallel and also substitute the epistemologically motivated metaphors of solid grounds. Whereas the former suggests stability and firmness, the latter suggests dynamic movement and transformation. I will argue that we must take seriously this switch in Husserl’s formulations and rethink the person and relation to its egoic kernel or core as essentially dynamic and developing.

But how can such endless totalities be understood? In his manuscripts, Husserl’s answer comes in two steps. These are sometimes presented as two aspects of one and the same process (e.g. Beyer 2012, Rinofner-Kreidl 2018), but I think that Husserl is actually describing two different ways or attitudes in which we can study the other person. On the one hand, we can study the other person – human or non-human – empirically as a psychophysical unity, but on the other hand, we can also study her as an ego with a comprehensive and dynamic ego-life. Both considerations have an empirical indication as their starting point (cf. Hua4, 276/288; Hua13, 444–445), but they proceed differently from this common point of origin, with two different thematic foci and two different sets of thematic positings, one focused on thingly objects and their causal and functional relations and the other focused on meanings and expressions and their mutual relations.<sup>5</sup>

So, an empirical apperception of the other, an experience of the other, is indispensable for both approaches, but whereas the former approach, the empirical, continues proceeding on this level, the latter approach, the intuitive-spiritual one, involves a change of attitude which eventually opens a dimension of depth. My experience of the other – of her behavior, glance, position or gesture – gives me, so to say, an “access” to her ego-life or refers me to this life. However, in order to study the other’s ego-life as such, as a life of an ego with its peculiar motivational nexuses and dynamic and open-ended structure, I must bracket the indicating reality, the behavior, and focus on the indicated alien ego-life. More precisely, my experiences of her ways of behaving function as indications of her ego-life, but when I have discovered this “access,” I have to suspend it and take as my theme her motivated life as such (Hua13, 445).

In the latter case I study the alien life as a variation of my own ego-life: It is given to me in empathy as an unrealizable modification of my own I-live, I-suffer, I-do, I-sense, I-feel, I-have-such-and-such-appearances, I-decide (Hua13, 455).

In the second volume of his *Ideas*, Husserl uses the term “intuition” in quotation marks (“*Intuition*”) to characterize the empirical understanding of others, and its holistic character, but he points out that a better term for this type of understanding would be “premonition” (*Vorahnung*) or “pre-seeing without seeing” (Hua 4, 286/273–274). The premonition of the other person is a special kind of ability to anticipate (and predict) the other’s behaviors and actions on the basis of our prior experience, including our experience of her as an individual and of other human beings in general (Hua4, 274/287). Such empirical understanding resembles genuine intuitive understanding (*Anschauung*) in that it is comprehensive and

---

154; cf. Bernet 1996; Mishara 1990; Depraz 2001; Bernet 2006; Smith 2010).

5 The fifth Cartesian Meditation makes clear that an indicative phenomenon is merely needed as a starting point of, or “access point” to, empathetic exchange. In order for a proper empathetic experience to be constituted, the experiencing subject has to capture the body over there as expressing the inner life of another subject and not merely as indicating someone’s existence (Hua1, 150/121, 151/124).

holistic. But it is not genuine insight since the actual motivational relations crucial to the other person remain obscure to me; they are not brought to full intuition (Hua4, 273/286) or clear insight by me (Hua4, 274/286), even if the general outline of the other's motivational nexuses is disclosed to me.

So, what is at issue is a “presentiment”, rather than a proper insight: “a pre-seeing without seeing, an obscure, more specifically symbolic, often ungraspably empty, premonition (*Voraus erfassen*)” (Hua4, 274/286).

The emptiness of this type of understanding means that the actual (*wirkliche*) motivational connections and nexuses of the other person's life are given to us merely as a *goal in anticipation*. They are determinate, however, in the sense that we can *follow* a clearly directed tendency and acquire a *chain* of actual (experiential) intuitions (Rinofner-Kreidl 2018). In so far as we have experiences of other people in general and of this person in particular, we can first grasp emptily the complex unity of the person's life, and then in the course of our own continuous experience of her we can fulfill this empty intending in insightful intuition, by means of the analysis of the *actual* nexuses of her motivational life, action and behavior (Hua4, 275/287). Husserl writes:

[T]his (...) premonition ought not to be confused with actual intuition. What we have here is the success of a more precisely determining apperception, which, like any apperception, offers a guiding line for the confirmation, in the course of experience, of the intentional nexuses, often extremely complex (Hua4, 275/187).

This type of understanding is processual and synthetizing; it proceeds step by step and by links and chains of experiences with mutual references. It consists of (i) an empty intention of the motivational nexuses that characterize and make up the other person's life, (ii) a guiding line (*Leitfaden*) based on experience of other people and this one in different situations, and (iii) the intuitive fulfilling of the intention by analysis of actual motivational nexuses.

We do not, and cannot, anticipate exactly, but we see the direction or know the “arsenal” of the other person's responses.

Husserl argues that even if empirical understanding of others often is mixed or confused with genuine intuitive understanding, these two modes of understanding must be kept conceptually separate.

[N]ot every judgment about the other is based solely on the, as it were, external style of his life, abstracted out of experience, for which I would not have to penetrate [*eindringen*] into the interior of her motivations and to represent [*vorstellen*] them in a fully lively way [*voll lebendig*]. I learn, however, to peer into the interiority of the other and to come to know inwardly the person herself, the motivational subject, which is precisely what bursts into view when I represent the other ego in the way it is motivated (Hua4, 273/285–286, translation modified).

Genuine or authentic understanding<sup>6</sup> is not a series of apperceptions but is a form of reflection that concerns the motivational structure of an alien ego-life and the ego as an inexhaustible source of activity. The actions of the other, when considered in this way, are not given to me

## 2. Towards Genuine Understanding of Others

---

6 In the first Intersubjectivity Volume, Husserl calls “authentic empathy” the genuine or authentic form of understanding of others in intuition.

as mental properties belonging or attached to some psychic or spiritual substrate; the whole model of substance is inadequate here. Rather the other ego is given as another spring or source of originary activity. In the first Intersubjectivity Volume, Husserl explains:

The relation between the self and its acts is specific; to be sure, it resembles the relation that a real substrate has to its features, thing and properties etc., but on the other hand it is very different. The self is not 'substance' in this sense, it is not an identical concrete that includes acts and states in itself as its features, that could provide an explication. Neither is the soul, the stream of the ego-life, such a substance [...] The self is no substratum, no 'carrier', but the self, a spring [*Quellpunkt*] [...]" (Hua13, 457)

In genuine understanding, the other is not posited as a psycho-physical reality but is posited as an ego with an ego-life, with her own ego-activities and passivities. To be sure, her ego-life is beyond my originary, direct intuitive reach, but, Husserl argues, I can go into a specific phantasmizing as-if mode of reflection and thus *accompany* the other in her actions and motivations. I can co-act with her, co-suffer her inclinations and co-decide with her, and so intuitively understand what she does and why she does so:

I not only empathize with his thinking, his feeling, and his action, but I must also *follow* him in them, his motivations becoming my quasi-motives, ones which, however, *motivate with insight* in the mode of intuitive fulfilling in empathy. I co-share in his temptations, I co-participate in his fallacies; in the "co-" there lies an inner co-living of motivating factors (...) that carry their necessity with them (Hua4, 275/287, translation modified).

In the Intersubjectivity Volumes, Husserl explains the nature of this quasi-life in greater detail, and points to a specific *duplication* of the self that it involves. He does this by comparing empathy to the relation that we have to our own past: We can return to our past life and go through our own motivations and actions in the past environment in an as-if mode. When we do this, we live in two different ways: we act *in the present* by inserting ourselves in the past and thus we, at the same time, as-if-feel, as-if-decide and as-if-operate *in the past*. Husserl's point about authentic empathy is that it has the same structure as this self-inspection: we can insert ourselves in the lives of others in the same way as we can insert ourselves in our own past lives. The first option is based on recollective de-presentation, and the latter on empathetic de-presentation:

Empathy as act is of course not living-in-my-past, but [it] is something similar: it is living-in-a-present or rather living-a-present-ego-life but in the modus of *quasi-life*, so that the self to which I empathize is not mine. And at the same time I am the actual, non-modified self that, transferring herself into the other, is living 'as if' [*gleichsam*] in the other (Hua13, 456; cf. Hua1, 145/115–116; Hua6, 189/186).

The as-if mode allows a special form of reflection: like non-modified feelings, decisions, and actions, also quasi-feelings, quasi-decision, and quasi-actions can be reflectively studied and evaluated. The other self is thus thematized in reflection as an alien ego-life, but it is not objectified in apperception:

[I]t allows a reflexion, a reflexion in respect to me, the non-modified, the empathizing self, and another reflexion in respect to the emphasized [other self] (Hua13, 456).

Husserl's explications of self-knowledge and self-understanding further illuminate the task of understanding others. The illumination comes from two arguments that we find in Husserl's exposition.

First, he distinguishes between two types of self-comprehension: mere empirical self-apperception that concerns myself as a unity of actual experiences and genuine self-understanding that also includes my *possibilities* of change and becoming over and above what I am and have been.

Second, Husserl argues that our understanding of alien motives is based on our self-understanding. In *Ideas*, he states: "[t]he underlying basis upon which is built the comprehensive motivated spiritual life of the other, and in conformity to which that life runs its course in an individual typical way, present themselves as 'variations' of my own" (Hua4, 275/288). In the first Intersubjectivity Volume, he writes: "Genuine empathy [...] is a modification of my own I-am and I-live, I-suffer-it, I-do-that, I-sense, I-feel, I-have-such-and-such-appearances, I-decide" (Hua13, 455). Thus, there is a crucial connection and analogy between self-understanding and the understanding of others.

We saw above that by introducing the concept of quasi-life, Husserl is able to argue that the proper, genuine understanding of others does not have to remain at the level of empirical apperception of their observed deeds and behaviors in comparison to other similar actions and similar actors.

But the parallel between self-understanding and the understanding of others suggest more than mere quasi-life. Since self-understanding essentially covers my *possibilities* of living and not just my actual life and its actual moments – experiences, motivations and actions –, also reflection on others can extend beyond the actual. Thus, it is not enough to study the other's actual motivations and deeds in the as-if-mode; we must also reflect on her possibilities of acting and being motivated, and so perform an *eidetic* reflection that is analogous to the one that we can perform in our own case. In the first Intersubjectivity Volume Husserl states: "The *absolute* self-experience (non-experiential, non-apperceptive) is the self-experience of the experiencing self. To it corresponds the absolute empathetic knowledge" (Hua13, 445).

The question then becomes how these foreign possibilities can be given in any intuitive way. I believe that we find an answer to this question by combining (i) Husserl's remarks about the quasi-motivations by which I can accompany the other's actual motivations and decisions for action and (ii) his discussion of the imaginary variation that allows me to study my own possibilities of being motivated.

My account here is based on the textual material that we find in the second volume of Husserl's *Ideas* and in his intersubjectivity manuscripts but the framing of my discussion is in Husserl's argument that ethical life is a life for *the best possible* in a person's life. This argument strongly suggests that reflection on human persons, on their actions and motivations, can and must – like self-reflection – extend beyond actualities and proceed in the sphere of possibilities. In Husserl's introductory lectures to ethics from the 1920s, we read:

[R]ather the moral self, the self of permanent and unceasing self-cultivation [*Selbsterziehung*], is the self that wants to make itself better, wants to change himself (as self) so that it can *eo ipso* [by this change] be only goodwill as an ethical self (Hua37, 162).

Mere empirical self-understanding is not sufficient for ethical self-shaping, nor for the proper understanding of other persons as open developing wholes. What is needed is an understanding of the motivational *possibilities* of the person in question.

Husserl describes a man who is known for his good heart but who in hurry acts without

### 3. The Task of Self-Understanding and its Implications to our Possibilities of Understanding Others

clear understanding of the condition of a person who appeals him for help. We tend to say: “He would have been charitable (...) if he had clearly understood the need of the one who appealed to him for help”. But this is far from proper consideration of the man in question, his character, for it is based on a superficial (contingent) truth about him, lacking consideration of his personal capacities, possibilities and potentialities. What is needed is a critical examination of the motives which are in operation as well as their alternatives, and all the intentional implications of these motives. The question is not just how he acts or would act in this or that situation, or who he is and has been thus far; the question is also about his possibilities of becoming *other* than he is and has been (Hua4, 252/264, 266/278).

In paragraph §60b in the second volume of *Ideas*, Husserl distinguished between empirical self-knowledge that is founded on the positing of the world and the self in the world, on the one hand, and the understanding of the motivational possibilities of oneself as a person, on the other hand. The former concerns the self as a posited unity of the manifold of position-taking acts and habitualities and faculties of acting; and the latter concerns the possibilities of the self and its capabilities of becoming, developing and changing.

*Empirical self-consciousness* is not limited to mere actualities. On the basis of my experience and knowledge of myself, I can also imagine myself in various situations and ask how my inclinations, affections and capabilities would direct my actions in these phantasized circumstances. This kind of knowledge is not confined to my actual deeds or performed actions and it is not a mere abstraction from them, but it also includes reflections on possible actions. But what remains constant, what is not varied, is the self as a subject of experiences and interests that is imaginatively inserted in different situations (Hua4, 265–266/278). Thus, this kind of self-comprehension is empirical in the sense that it is founded on the positing of the self as an already constituted unity of experiences and does not question or neutralize this posited whole. Husserl explains:

If I now phantasize, if I settle myself (as the one I am) into a phantasized actuality or into the world given in the neutrality modification, into the familiar world transformed in phantasy some way or other, then I am judging how such and such motives (more precisely: the quasi-motives of this phantasized environment) *would* affect me; how I, as the one I am, would act and could act, how I could, and how I could not, judge, value, and will. In that way I judge, or I can judge, empirically, *on the basis of my experiential knowledge of myself*, with respect to myself, with respect to *the ego constituted for me in empirical apperception as experiential ego* (Hua4, 265–266/278, emphasis added).

But there is another mode of self-reflection that is not, and cannot be, based on previous experiences of oneself since it concerns oneself in the constant process of becoming and developing. Husserl states:

Thus, the substrata of motivation, the orientations and the powers of the motives are different. How do I come to know them? I do so, as the one I am, by means of phantasizing presentification of possible situations, in which I ‘reflect’ on what kind of sensuous or spiritual stimulation would affect me [...] Upon reflection I could [...] say that I, as I used to be, [...] would not have been able to do something or other. But at present I can act in that way and would do so. When I say this, I am not basing myself on experience but on the fact that I can test my motives at the very outset (Hua4, 267/279–280).

In this context, Husserl emphasizes that the unity of a person is fundamentally different from the unity of a thing (Hua4, 266/278–279, 274/286–287). Unlike a thing, a person can change in

radical ways without losing her identity as the ego subject of the original faculties and powers (*Urvermögen*) of willing, valuing, and thinking. The whole complex of motivations can change as well as their directions and powers (Hua4, 254–255/267, cf. Hua27, 30–32; Husserl [1923] 1997, 207–208). Latent motivations can emerge and potent motivations may retreat (Hua4, 252/264; cf. Hua6, 240/237–238).

The difference between persons and things includes three crucial and related structural features.

First, the thing is predictable; it cannot really surprise us by disclosing completely unexpected aspects. Husserl writes: “In the apprehension of a thing there is nothing that, in principle, is new. If there were, then it would already be the start of the constitution of a new stratum of unity” (Hua4, 266/278). The person, on the contrary, can always act and react differently than expected without losing her identity – not just in the other’s eyes but also in her own.

Second, things come in kinds but persons are given to us as unique individuals. Husserl explains: “Each thing is of a certain kind. If one knows the kind, the rest can be dispensed with” (Hua4, 274/286). The person, on the contrary, is an individual ego of intentional activities and passivities and of position takings and can actively change the course of her own life (Hua4, 277–278/289–290). Even more: the ego, and only the ego, is the unit that is absolutely individuated (Hua4, 299–302/313–316; cf. Heinämaa 2018).

Finally, a person is a dynamically developing unity (Hua4, 251–253/263–265; cf. Hua3, 132–141/184–193). In the third volume of *Ideas*, Husserl emphasizes the fact that persons as spiritual units differ from things primarily in that their moments and states are non-recurring (Hua5, 17, cf. 44). Moreover, Husserl’s concepts of motivation and willing imply the idea of a motivational change, that is the idea that we can be motivated to abandon, not just our formed ways of acting, but also our former ways of being motivated to act and thus become different in our motivational nexuses. Thus, position in a historical setting or membership in a community – professional, social, or natural – does not limit or exhaust the person’s possibilities of change and development. Husserl explains:

Originally, I am actually not a unity of associative and active experiences (if experience means the same as it does in the case of the thing). I am the subject of my life, and the subject develops in living [...] The ego does not originally stem from experience – in the sense of an associative apperception in which are constituted unities of manifolds of a nexus – but from life (it is what it is not *for* the ego, but it is itself the ego) (Hua4, 252/264, translation modified).

I have argued that if knowledge of self and understanding of others is to serve ethical reflections, and ethical self-shaping and ethical judgment of others, it must not be confined to the actual but must involve reflection on the other’s possibilities. It concerns the other not just as an actual agent in apperception, anticipation and extrapolation but also as a possible agent. Understanding of the other person means that I – the one who aims at understanding – live through her life in the modality of as-if, I live through her actions and motivations in the quasi-mode. But for ethical reflection to be possible, quasi-life must cover not just the other’s actual, past and present, actions and motivations but also her possibilities of becoming different and renewing herself. The analogy is not just between myself as I am and have become at the present and the other as an actual person but is between two individual essences: myself as I can be, as I can act and be motivated, and the other as she can be.

#### 4. Summary

### BIBLIOGRAPHY

- Bernet, R. (1996). The unconscious between representation and drive: Freud, Husserl and Schopenhauer. In John J. Drummond & J. G. Hart (eds.), *Truthful and Good: Essays in the Honor of Robert Sokolowski*. Dordrecht, London, Boston: Kluwer, 81–95;
- Bernet, R. (2006). Zur Phänomenologie von Trieb und Lust bei Husserl. In D. Lohmar & D. Drik Fonfara (eds.), *Interdisziplinäre Perspektiven der Phänomenologie: Neue Felder der Kooperation: Cognitive Science, Neurowissenschaft, Psychologie, Soziologie, Politikwissenschaft und Religionswissenschaft*. Dordrecht: Springer, 38–53;
- Beyer, C. (2012). Husserl on understanding persons. In C. Fricke & D. Føllesdal (eds.), *Intersubjectivity and Objectivity in Adam Smith and Edmund Husserl*. Berlin: De Gruyter, 93–116;
- Crespo, M. (2015). Husserl on personal aspects of moral normativity. *Ethical Perspectives*, 22(4), 699–722;
- Depraz, N. (2001). Pulsion, instinct, désir: Que signifie Trieb chez Husserl. *Alter*, 9, 113–125;
- Donohoe, J. (2010). The vocation of motherhood: Husserl and feminist ethics. *Continental Philosophy Review*, 43(1), 127–140;
- Hart, J. (2006). The absolute ought and the unique subject. *Husserl Studies*, 22(3), 223–240;
- Heinämaa, S. (2014). Husserl's ethics of renewal: A personalistic approach. In M. Tuominen, S. Heinämaa & V. Mäkinen (eds.), *New Perspectives to Aristotelianism and Its Critics*. Leiden, Boston: Brill, 196–212;
- Heinämaa, S. (2018). Embodiment and bodily becoming. In D. Zahavi (ed.), *The Oxford Handbook of the History of Phenomenology*. Oxford: Oxford University Press, 533–557;
- Husserl, E. (Hua1). *Cartesiansche Meditationen und pariser Vorträge, Husserliana I*. Ed. Stephan Strasser. Den Haag: Martinus Nijhoff, 1950. In English *Cartesian Meditations*, trans. Dorion Cairns. The Hague: Martinus Nijhoff, 1960;
- Husserl, E. (Hua3). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, Erstes Buch: Allgemeine Einführung in die reine Phänomenologie, Husserliana III*. Ed. Walter Biemel. Den Haag: Martinus Nijhoff, 1913. In English *Ideas: General Introduction to Pure Phenomenology*, trans. W.R. Boyce Gibson, New York, London: Collier, 1962;
- Husserl, E. (Hua4). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, Zweites Buch: Phänomenologische Untersuchungen zur Konstitution, Husserliana IV*. Ed. Marly Bimel. Den Haag: Martinus Nijhoff, 1952. In English *Ideas Pertaining to a Pure Phenomenology and to a Phenomenological Philosophy, Second Book: Studies in the Phenomenological Constitution*, trans. R. Rojcewicz and A. Schuwer. Dordrecht, Boston, London: Kluwer Academic Publishers, 1993;
- Husserl, E. (Hua5). *Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, Drittes Buch: Die Phänomenologie und die Fundamente der Wissenschaften, Husserliana V*. Ed. Marly Biemel. Haag: Martinus Nijhoff, 1952;
- Husserl, E. (Hua6). *Die Krisis der europäischen Wissenschaften und die transzendente Phänomenologie: Eine Einleitung in die phänomenologischen Philosophie, Husserliana VI*. Ed. Walter Biemel. Den Haag: Martinus Nijhoff, 1954. In English *The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy*, trans. David Carr. Evanston: Northwestern University, 1988;
- Husserl, E. (Hua11). *Analysen zur passiven Synthesis, Aus Vorlesungs- und Forschungsmanuskripten, 1918–1926, Husserliana XI*. Ed. Margot Fleischer. Den Haag: Martinus Nijhoff, 1966. In English *Analyses Concerning Passive and Active Synthesis: Lectures on Transcendental Logic, Collected Works, Volume IX*, trans. Anthony J. Steinbock. Dordrecht: Springer, 2001;
- Husserl, E. (Hua13). *Zur Phänomenologie der Intersubjektivität: Texte aus dem Nachlass, Erster Teil: 1905–1920, Husserliana XIII*. Ed. Iso Kern. Den Haag: Martinus Nijhoff, 1973;
- Husserl, E. (Hua15). *Zur Phänomenologie der Intersubjektivität: Texte aus dem Nachlass, Dritter Teil: 1929–1935, Husserliana XIII*. Ed. Iso Kern. Den Haag: Martinus Nijhoff, 1973;

- Husserl, E. (Hua27). *Aufsätze und Vorträge 1922–1937, Husserliana XXVII*. Eds. Tom Nenon & Hans R. Sepp. Den Haag: Kluwer Academic Publishers, 1988;
- Husserl, E. (Hua37). *Einleitung in die Ethik, Vorlesungen Sommersemester 1920 und 1924, Husserliana XXXVII*. Ed. Henning Peucker. Dordrecht: Kluwer Academic Publishers, 2004;
- Husserl, E. [1923] 1997. Wert des Lebens. Wert der Welt. Sittlichkeit (Tugend) und Glückseligkeit <February 1923>. *Husserl Studies*, 13(3), 201–235;
- Melle, U. (2007). Husserl's personalistic ethics. *Husserl Studies*, 23(1), 1–15;
- Mishara, A. (1990). Husserl and Freud: Time, memory and the unconscious. *Husserl Studies*, 7, 29–58;
- Rinofner-Kreidl, S. (2018). Phenomenological intuitionism and its psychiatric impact. In Fuchs, T., Breyer, T. & C. Mundt (eds.), *Karl Jasper's Philosophy and Psychopathology*. Dordrecht: Springer, 30–60;
- Siles i Borràs, J. (2011). *The Ethics of Husserl's Phenomenology*. London, Oxford, New York: Bloomsbury;
- Smith, N. (2010). *Towards Phenomenology of Repression – a Husserlian Reply to the Freudian Challenge*. Acta Universitatis Stockholmsiensis 34. University of Stockholm;
- Trincia, F. S. (2007). The ethical imperative in Edmund Husserl. *Husserl Studies*, 23(3), 269–186.





---

ROBERTA LANFREDINI  
*University of Florence*  
*lanfredini@unifi.it*

---

# PHENOMENOLOGICAL EMPIRICISM

---

## *abstract*

*The aim of this paper is to compare Phenomenological Empiricism with two different kinds of Empiricism: Classical Empiricism (represented by Hume) and Logical Empiricism (represented by Schlick). Phenomenological Empiricism is at the same time a radical and sophisticated theory of experience, in which intentionality, ideation, material a priori, a complex notion of data and of intuition play a crucial role.*

---

## *keywords*

*phenomenology, empiricism, intentionality, ideation*

## 1. Radical Empiricism

The aim of this paper is to show how Phenomenology is a radical and at the same time complex (anti-reductionist) Empiricism. To this end, I will underline similarities and differences between Phenomenological Empiricism and other forms of Empiricism. I will consider in particular two important kinds of empiricism: Classical Empiricism and Logical Empiricism. A third and very important type of empiricism is the so-called Neutral Empiricism (that of James, Mach and also, I believe, Bergson). This paper will be limited in scope to the examination of the first two kinds of empiricism.<sup>1</sup>

First of all, two preliminary questions. What do we mean by the expression “radical empiricism”? And in what ways can Phenomenology be considered a radical (even if sophisticated) kind of empiricism?

In a radical empiricism it all begins and ends in the field of Experience. Husserl says:

It must always be borne in mind here that *whatever physical things are* – the only physical things about which we can make statements, the only ones about the being or non-being, the being-thus or being-otherwise of which we can disagree and make rational decisions – *they are as experienceable physical things*. It is experience alone that prescribes their sense; and, since we are speaking of physical things in fact, it is actual experience alone which does so in its definitely ordered experiential concatenations.<sup>2</sup>

In a radical empiricism, being misled or deceiving oneself also falls within the range of phenomena.

As Husserl says:

I may be deceived as to the existence of the object of perception, but not as to the fact that I do perceived it as determined in this or that way, that my percept’s target is not some totally different object, a pine-tree, e.g., instead of a cockchafer. The self-evidence in characterizing description (or in identification and distinction of intentional objects) has, no doubt, its understandable limits, but it is true and genuine self-evidence.<sup>3</sup>

---

1 For the relationships between phenomenology and neutral pragmatism see Lanfredini (2017).

2 See Husserl, 1983, p. 86.

3 Husserl, 2001, p. 297.

Husserl's position, however, should not be confused with the Cartesian *videre videor*. According to Descartes, I could be deceived about what I see, but not about the fact that I am seeing it now. More radically, according to the so called static version of phenomenological empiricism, it is inappropriate to talk about deception not only from a noetic point of view (seeing something) but also from a noematic point of view (what I see).

I think that, against the so-called West Coast Interpretation<sup>4</sup> (and in favour of Gurwitsch<sup>5</sup>), Husserl never proposed a realistic design of intentionality in a metaphysical sense. The principle by which objects that are not intended are not given is a principle of phenomenology with no exceptions. In a descriptive context this refers us back to the impossibility of making an object-based analysis without connecting it with some form of experience. In other words, in the phenomenological object there is always contained (from an ontological point of view) the reference to a state of consciousness: the red I detect necessarily leads back to a determined sensation of redness; the chair I see necessarily leads back to a determined perception of that chair; the thought property necessarily leads back to a determined modality of thought, and so on.

Any phenomenon, in order to be such (that is, to become manifest) necessarily leads back to an intentional *Erlebnis*. The necessity, contained in the structure of the phenomenon, of the reference to a state of consciousness allows the limit of Phenomenological Ontology to be defined.

This is one of the main differences with respect to Hume's empiricism. The phenomenological notion of data includes directness but not simplicity, even in fundamental elements such as sensations. The phenomenological notion of data goes beyond the amorphous and undiversified notion of *idea* to embrace a bi-polar concept of experience composed by a noetic pole and a noematic pole. According to Hume, and also Berkeley and Locke, a distinction between feeling and felt property, hyletic data and objectual data does not exist. According to Husserl, on the contrary, this distinction is absolutely fundamental.

The fact that for Husserl (*contra* Hume) the intended object may not effectively be contained within the state of consciousness (the singer's song, says Husserl, is distinct from the act of singing), while still being essentially bound to it (there can be no song whose singer can dispense with the act of singing), entails a distinction between two ontological dimensions: the dimension relating to subjective or noetic object (states of consciousness) and the dimension relating to real, objectual, or intended object.

These environments delimit two distinct but not separated regions: the first region, the noetic, which consists of adequate, non-perspectival objects (those given in a complete manner and "without residues", so to speak - eg. seeing a red spot, hearing the sound of violin, imaginings the god Jupiter); the second region, the noematic, which consists of objects that are inadequate, incomplete, essential perspectival and unilateral (eg. the red spot that is seen, the sound of the violin, the imagined god Jupiter). This distinction isolates domains of objects which, although ontologically distinct, are nevertheless connected by an essential and necessary relationship of unilateral foundation<sup>6</sup>. In this sense, phenomenological analysis (at least from a static point of view) is technically correlative.

In Husserl's opinion, Hume had passed right through phenomenology with his eyes closed. The first blindness consists in its denial of transcendence. Phenomenological transcendence is not

## 2. Hume's blindness

---

4 See Føllesdal, 1969, 1982; Dreyfus, 1979, McCulloch, 2003.

5 See Gurwitsch, 1967.

6 See Husserl, 2001, §16.

an absolute but a relative one: that is, relative to the state of consciousness that constitutes it. The expression “phenomenological transcendence” consists of two essential conditions: first, the intended object is not effectively contained within the act which intends it; and second, the intended object refers in an essential manner to an act that intends it.

The first condition confirms the *reality* of the object, the second denies its absolute, or metaphysical, reality. So, existence is not a property of the object in itself, but a property of the act that intends the object: when we say that an object exists we are in reality asserting that the object is understood *as existing*; that is, we are using a specific modality, a *thetic* or positioning character. In perception, as in memory, unlike in fantasy or in imagination, the object appears *as existing*. Conceived in this way, “authentic” perception, so to speak, is no different from hallucination, as far as content or description are concerned. Indeed, hallucinatory perception benefits from a positioning character just as much as other perceptions do, and as a consequence the hallucinated object appears to us *in flesh and blood* just like any other perceived object (contrary, for example, to what is maintained in Searle).<sup>7</sup> The “reality” of the authentic object of perception lies in a sort of operative and contextual attitude. So, the materiality of things depends upon the circumstances and the context in which they are placed. If we consider a thing by itself, distinguishing between something and its phantom – i.e. its pale, empty, and ghostly counterpart – becomes virtually impossible. The ghost of a certain thing has all the essential features that render that thing exactly what it is and nothing else: essential features that are dispersed throughout an extension. In this sense, we would see rainbows and blue skies but we could not define them as material things. On the other hand, if we consider the thing within a given context, the thing and the ghost of that thing cannot be regarded as the same element. Things exist, are real, substantial, and causal (these terms are all synonyms) when they behave in a certain way. In this sense, real (or material) properties are, *ipso facto*, functional links: for example, causal links. In order to get to know the reality of a given thing, we must be able to predict its behavior under a certain force, pressure, when it is smashed up, cooled down, heated up and so on. In the multiplicity of its dependence relations the real thing will retain its own identity.

### 3. A world outside our world

The possibility of the thing in itself turns out, in conclusion, to be decisively denied by Husserl exactly as by Hume. An object that cannot be intended in principle is in fact an effective absurdity. But Husserl’s criticism is more sophisticated than Hume’s. Husserl admits that a thing in itself is not a matter of non-sense, hence of a contradiction. Asserting the existence of a thing in itself does not in fact mean violating any formal law (such as seems to happen in cases of syntactical nonsense such as “the tree is and” or semantic nonsense such as “round square”).

The hypothetical assumption of something real outside this world is, of course, logically possible; obviously it involves no formal contradiction. (...) When that is taken into account the formal-logical possibility of realities outside the world, the *one* spatiotemporal world, which is *fixed* by our *actual* experience, materially proves to be a countersense.<sup>8</sup>

On the other hand, the absurdity of the thing in itself cannot be ascribed to material countersense (that violation which we encounter in examples such as “color without extension”).

---

<sup>7</sup> See Searle, 2012.

<sup>8</sup> Husserl, 1983, pp.108-9.

The absurdity of the thing in itself seems to reside in a third order of reasons, which we could call motivational, according to which, in order for an object to be given, it must exhibit a rootedness in an actual experience. And this is the case for the thing in itself, which turns out to be impossible not because formally nonsensical, nor because it is materially counter-sense, but because it does not have in principle a bond of motivational synthesis with an actual experience. In the absence of such a chain of connections in experience, the object vanishes. A world outside our world is a concrete absurdity. There are two obligations which the notion of world must satisfy so as not to fall into concrete absurdity. The first is that any given phenomenon must effectively contain within it the reference to an experience. The second, which constitutes the foundation of so-called genetic phenomenology, is that any phenomenology, in order to be configured as such, must exhibit an original nexus (however remote) with an actual experience, with the acknowledgement of an object in flesh and blood. The thing in itself violates both these conditions (the first formal, the second material): on the one hand it is a thing which in principle severs the intentional nexus (which any object must satisfy in order to be declared as such); on the other it negates that rootedness in a concrete experience.

In conclusion: making reference to a world *outside* our world, hence to the thing in itself, does not entail either the violation of the purely material ontological law according to which two types which are founded one on the other are in fact separate (as in the case of color and extension), nor the violation of the purely formal ontological law by which two disconnected species cannot live within the same singularity (as in the case of the square and the round). In the thing in itself, we are in fact witnessing the violation of a further, still more fundamental law, the law by which the notion of thing necessarily contains the reference to lived experience. If we conceive the two objects of object and experience as unconnected categories linked by a relationship of unilateral foundation, the expression “thing in itself” expresses a counter-sense that is both formal and material. We can talk about constitutive and concrete absurdity. The thing in itself is a constitutive (formal) absurdity because it makes reference to something that contains the reference to a constitutive function and at the same time removes itself from it. Therefore it is not possible, for constitutive-formal reasons, to refer coherently to a thing in itself, simply because the thing in itself is not something, or, which is the same, it is a non-thing. Although not senseless, the thing in itself is impossible since it violates the formal law by which the notion of thing, containing *in itself* the reference to a corresponding intention, cannot at the same time free itself from such a reference. For this reason it turns out, though certainly in a different acceptation from syntactical and semantic senselessness, to be unthinkable and not merely imperceptible and unimaginable, as happens in purely formal and purely material counter-sense.

The thing in itself is at the same time a concrete (material) absurdity because it is not rooted in any original and lived experience. Therefore, it is not possible, for material reasons, to refer consistently to a thing in itself, because it is not available.

Hume's second blindness rests in his negation of the essential, eidetic dimension of experience. As is well known, Husserl's theory supports an ideative notion of data, against the empiricist theory of abstraction (see in particular the attention theory (Mill), the representational theory (Locke, Berkeley) and the theory of *distinctio rationis* (Hume)). At the same time Husserl claims a material and not formal conception of data (*contra* Cassirer). I think that the ideative notion of data not only is not in contradiction with phenomenological empiricism, but completely corroborates it. The event, being invariant, requires ideation; for example, invariant to variations of lighting conditions, changes in positions, and so on. Data, being invariant, are not without structure. Phenomenological ideation has the following characteristics. First, it is not

#### **4. Form and content of the experience**

an abstraction: it does not function by means of exclusion or negation of some characteristics for the benefit of another and it is not founded on any recording of similarities (*contra* Hume and Classical Empiricism). Second, it is not a concept: it does not function by means of categorization. There is no need of a good theory of concept constitution for a good theory of data (*contra* Kant). Third, it is not formal: it is not a functional but a material process: that is, a process founded on specific features of experience (*contra* Cassirer).

That last statement introduces the second point that I intend to address: the relationships between Phenomenological Empiricism and Logical Empiricism, respectively an empiricism with and without intuition.<sup>9</sup>

A crucial question for Schlick, as well as for Logical empiricism in general, is the total insignificance of the concept of content (in particular, intuitive content) for both a theory and a practice of knowledge. Besides, it is this thesis (related to a notable distinction between *Kennen* and *Erkennen*) which marks – on Schlick’s explicit admission – his radical divergence from Husserl.

According to Schlick the content is inexpressible and indescribable: that is, linguistically untouchable: “the difference between structure and material, between form and content is, roughly speaking, the difference between that which can be expressed and that which cannot be expressed”.<sup>10</sup>

Let’s consider the famous example, in Schlick, of a person who is born blind. This example is perfectly analogous to Husserl’s example (contained in *The Idea of Phenomenology*<sup>11</sup>) of the deaf man. Let’s consider the perception of green and look at the ineffable quality of green which makes the essence of the content. This quality is accessible only to beings endowed with eyesight and power of color perception; it couldn’t possibly be conveyed to a person born blind. Shall we conclude, Schlick asks, that such a person could not understand any of our statements about color, that they must be quite meaningless to him because he can never possess the green content? Schlick’s answer to this question is no. We can communicate to the person born blind, as we can to a person who can see, the meaning of “green”. Nevertheless what we communicate is not the content – the *greenness*: “since content is incommunicable by language, it cannot be conveyed to a seeing man any more or any better than to a blind one”.<sup>12</sup> What can be communicated (or, which is for Schlick the same, expressed) is the fact that something exists – which we call green – and that it is something possessing a certain structure or belonging to a certain system of internal relationships.

I can give a particular description of this green leaf lying on my desk by placing the color in a certain order. I assert, for instance, that is a bright green, or a rich green, or a bluish green, trying to describe the green by compare it to other colors. Evidently it belongs to the intrinsic nature of our green that it occupies a definite position in a range of colors and in a scale of brightness, and this position is determined by relations of similarity and dissimilarity to the other elements, in this case shades, of the whole system. In this sense, quality, every quality, has a certain definite logical structure: “in this way every quality (for instance, the qualities of sensation: sound, smell, heat, etc., as well as color) is interconnected with all others by internal relations which determine its place in the system of qualities”.<sup>13</sup>

In such a perspective the difference between a color-blind person and a normal one is existentially relevant but philosophically limited. Both *are knowing* the same thing. And both

---

<sup>9</sup> I have addressed this issue in more details in Lanfredini, 2003.

<sup>10</sup> Schlick, 1938, p. 291.

<sup>11</sup> Husserl, 1966.

<sup>12</sup> Husserl, 1966, p. 295.

<sup>13</sup> Husserl, 1966, p. 294.

*are not knowing* the same thing. Both know the structure in which the green is placed and its position in the color spectrum. But, according to Schlick, both do not know the content of the green. The only difference between the two consists in their ability to enjoy (or not enjoy) the green color. This is absolutely important as far as actual life is concerned, but just as absolutely irrelevant for the knowledge of color.

Schlick says:

I can perceive a green leaf; I say that I perceived it if (among other things) the content “green” is there, but it would be nonsense to say that I perceive this content.<sup>14</sup>

The content, in the sense of intuitive content, is simply there. The verb used by Schlick in this regard is “enjoying”, the nearest equivalent to the German “*Erleben*”.

Schlick says again:

Here we uncover the great error committed by the philosophy of intuition: the confusion of acquaintance (*Kennen*) with knowledge (*Erkennen*).<sup>15</sup>

So, Schlick claims a sort of epistemological (although not-ontological) eliminativism, very similar to that held by Churchland, according to which the qualitative element is irreducibly outside of an adequate theory of knowledge. When I look at the blue sky and lose myself entirely in the contemplation of it without thinking, then I am enjoying the blue, I am in a state of pure intuition. The blue fills my consciousness completely. But that does not mean knowing what blue really is. The Meaning of the word “blue” is entirely included in the structure of the intuitive content.

In *The idea of Phenomenology* Husserl’s example is perfectly analogous to Schlick’s, with respect to a person born deaf. A person born deaf knows that sounds exist, and sounds make harmonies. But he cannot understand how sounds do this, how symphonies are possible. He cannot represent a thing of that kind, he cannot perceive it; and perceiving it he can’t understand *the how*. No kind of physical or psychological theory about color can add anything to this “pure vision” which, according to Husserl, establishes the sense of color.

This thesis presents, from a phenomenological point of view, at least two problems. First, the necessity of an intuitive content for the determination of the empirical knowledge. Second, the necessity of an intuitive content for the configuration of the same structure.

Without intuition it is not possible to intend determined objects. According to Husserl, determination comes with the fulfillment of an empty intention. Without an intuitive act, by which I mean, a fulfilled act, it is not appropriate to speak about knowledge. So, if confronted with a brilliant green leaf, a sighted person and a blind person share the same thing, an empty intention aiming at a “pure something”.

But in no case can a blind person believe himself to know either the leaf or its particular green. Knowing something means necessarily to have a fulfilled act, that is an act equipped with an intuitive content. Without a full, intuitive content every determined knowledge is impossible. And without a full, intuitive content the structure of green (and the structure of colors in general) is inscrutable. According to the phenomenological point of view, intuition is a necessary condition for knowing something because it is a necessary condition to have

## 5. Content and determination

---

<sup>14</sup> Schlick, 1974, p. 319.

<sup>15</sup> Schlick, 1974, p. 83.



a determined intention of something. Knowing the structure of something requires a prior experience, and not vice-versa.

In fact, the rigorous separation between Schlick and Husserl can be analyzed further. The problem, I think, is to establish what we mean by intuitive content, a concept that has a more complex structure than Schlick has supposed. But as we have seen, the concept of living does not exhaust the notion of intuition. The distinction between *experiencing* (*erleben*) the content and *apprehending* (*auffassen*) or perceiving a property or an object makes Schlick's treatment more complex. We experience (or enjoy, as Schlick says) acoustic sensations, but we hearing (we perceive) the singer's song. So, contrary to what Schlick says ("when I gaze at a red surface, the red is part of the content of my consciousness"<sup>16</sup>), red is not a part of the content of my consciousness. Affirming that content is simply there and simply present means advocating a simplistic and undifferentiated vision of the concept of intuition. We experience content, but we mean objects that go beyond these contents in the strict sense.

Another clarification: for Schlick, content (in the sense of intuitive content) has nothing to do with knowledge. There is still more: "the most fundamental mistake of philosophy of all times"<sup>17</sup> is to identify knowledge with immediate awareness or with intuition. But Husserl has never identified knowledge and intuition, even if he considers intuition as necessary condition for knowledge. When I hear a sound or see a color, it is not with these acts of hearing and seeing that I come to understand what a sound or a color is. Knowing and having acquaintance with are not at all synonymous even for Husserl. According to Schlick, knowledge is a result of an act of comparing, recognizing, naming. According to Husserl, knowledge means to recognize, to identify the intended object, too. Nevertheless, this does not mean, as Schlick suggests, to differentiate two types of knowledge: one conceptual and one intuitive. On the contrary, knowledge is the result of an integration between two components. This is exactly the reasoning which makes it possible to in phenomenological terms to distinguish between *discrimination* of something (the result of a simple act of perception) and *identification* of something as such and such (the result of the conjunction of a meaning act and a perceptive act). This important distinction does not seem present in Schlick's perspective.

## 6. Phenomenal Intentionality

Husserl's analysis is more subtle in comparison with that of Schlick. In fact, he distinguishes sensation from perception on the one hand, and perception from conceptual elaboration on the other. Perception, for Husserl, is not a judgment; it is not knowing something. The distinction marks the important difference between an epistemic conception and a non-epistemic conception of perception. Husserl accepts the second one; in this sense he corrects Kant's famous principle according to which "intuition without conception is blind, and conception without intuition is empty". While the second part of the affirmation is without doubt true, the first needs a correction: it is not the concept that allows intuition to see; perception has already, in itself, an organization, a structure which does not have any reference to conceptualization.

Then, according to Husserl, it is not necessary that our consciousness experience acquires propositional contents. In particular, conscious experiences can represent something in a certain way regardless of inferential or propositional content: that is, regardless of beliefs and concepts. Conscious experiences are intrinsically related to a qualitative format (*qualia*). Qualitative states are independent of their descriptive value, so that we can have a conscious experience without having a description of something.

---

<sup>16</sup> Schlick, 1938, p. 102.

<sup>17</sup> Schlick, 1938, p. 318.

There is in fact an important distinction between perceptual discrimination (non-propositional) and perceptual identification (propositional).

Against the so called *tracking theories* (TT)<sup>18</sup> whose main endeavor is to naturalize the content of qualitative states by reducing it to tracking relations holding between the phenomenality of consciousness and physical properties of environment, Husserl would endorse the so called *phenomenal intentionality research program* (PIT)<sup>19</sup>. According to PIT, intentionality is basically a matter of phenomenal consciousness: any qualitative state has a phenomenal property which expresses a specific and determined qualitative state. Furthermore, it is precisely by virtue of this richness of phenomenal properties that it is impossible to think about different experiences with different propositional content. The propositional content depends on phenomenal properties, and not vice versa. If the qualitative spectrum of our experiences were not characterized by a plurality of phenomenal properties, it could not be available to a propositional format.

Accordingly, the experience's propositional content depends in a strong sense on having a plurality of phenomenal properties that characterize the experience in itself.<sup>20</sup>

In conclusion, the famous distinction between *Kennen* and *Erkennen* is not at all denied by Husserl. Certainly for Husserl, as for Schlick, *Kennen* is not knowledge in a proper sense. But knowledge needs the *Kennen* in order to acquire determined knowledge. So, there is a crucial difference between the two perspectives. Schlick declares explicitly that "intuition and conceptual knowledge do not both strive for the same goal; they move in opposite directions"<sup>21</sup>. According to Husserl, in contrast, intuition is an integrating part of the genuine process of knowledge. Authentic knowledge cannot function without a qualitative factor. To determine and consequently to identify a red object, we must have an experience of "what it is like" (using Nagel's expression) to see something red. We have knowledge when we have an acknowledgment. But acknowledgment necessarily implies an act of intuition.

So, I think that the main element of the deep disagreement between Husserl's empiricism and Schlick's empiricism resides in a different philosophical conception of the notion of intuition (and, thus, in the notion of experience). Schlick speaks about intuition as "an exceptionally close relation between subject and object", something mystical and completely inexpressible (the borderline case of this is the relation between consciousness and God). But this is a mere caricature of the concept of intuition. Intuition has an internal structure.

The second difference between Schlick and Husserl resides in the different value that the two attribute to the concept of intuition. For Schlick, intuition is radically outside the dimension of knowledge; according to Husserl, in contrast, intuition is an integral part of knowledge. Without it, no determined knowledge is possible. In this sense, I think that the content empiricism advocated by Husserl is more radical than the formal empiricism advocated by Schlick.

In conclusion, phenomenology is a kind of radical empiricism. But it is not a reductionist form of empiricism. Husserl's position has not assimilated to classical Empiricism for these reasons. First, the phenomenological notion of data includes directness and immediacy but not simplicity. In Classical Empiricism, consciousness is a place, and the objects are immanent data: that is, sensations, ideas, perceptions, and so on. For Husserl, consciousness

## 7. Conclusions

---

18 Dretske, 1981, 1988, 1996, 2003; Millikan, 2009; Papineau, 1987.

19 Kriegel, 2008, 2013.

20 See Zipoli Caiani, 2019.

21 Schlick, 1938, p. 82.

is, on the contrary, not a closed place, and the object is not inside consciousness, as in a sack. Second, the phenomenological notion of data includes reality, even if *relative*, or, if you will, empirical. The phenomenological notion of reality is functional, relational, operative and not substantial in any sense of the word. Third, the phenomenological notion of absolute reality is an absurdity (as in Hume), but the argument in Husserl is much more sophisticated and more fully articulated. Husserl distinguishes different forms of absurdity: nonsense, counter-sense and constitutive-motivational (or concrete) absurdity. A *reality in itself* reflects the last kind of absurdity. Finally, the phenomenological notion of data requires an ideative process that fixes an invariant to the variations of different conditions. Husserl's position is profoundly different from Logical empiricism for these reasons: a) the phenomenological notion of intuition is more elaborated and sophisticated by comparison with that of Schlick. For example, in Husserl an important difference exists between discrimination of something and identification of something. Both imply intuition, but in a different sense; b) the phenomenological notion of intuition is absolutely essential for knowing something. Without intuitive, qualitative content, the empirical knowledge is lacking in determination. Moreover, without intuitive, qualitative content we cannot set a structure of something; for instance the structure of color, the structure of sound, the structure of physical things, and so on.

#### REFERENCES

- Dretske, F. (1981). *Knowledge and the Flow of Information*. Boston: MIT Press;
- Dretske, F. (1988). *Explaining Behavior: Reasons in a World of Causes*. Boston: MIT Press;
- Dretske, F. (1996). *Naturalizing the Mind*, Boston: MIT Press;
- Dretske, F. (2003). Experience as Representation. *Philosophical Issues*, 13(1), 67-82;
- Dreyfus, H. (1972). Husserl's Perceptual noema. In L. Embree (Ed.), *Life-World and Consciousness: Essay for Aron Gurwitsch*, Evanston: Northwestern University Press, pp. 135-17;
- Føllesdal, D. (1969). Husserl's notion of noema. *Journal of Philosophy*, 66(20), 680-7;
- Føllesdal, D. (1982). Hintikka on Phenomenology. In E.Randall, E.Auxier, L.E.Hahn (eds.), *The Philosophy of Jaakko Hintikka*, Library of Living Philosophers, vol. XXX, Open Court, Illinois, 73-387;
- Gurwitsch, A. (1967). Husserl's Theory of the Intentionality of Consciousness in Historical Perspective. In E.N. Lee-Mandelbaum (Ed.), *Phenomenology and Existentialism*. Baltimore: John Hopkins University Press, 46-57;
- Kriegel, U. (2008). Real Narrow Content. *Mind and Language*, 23(3), 304-328;
- Kriegel, U. (2013). The Phenomenal Intentionality Research Program. In U. Kriegel (Ed.), *Phenomenal Intentionality*. Oxford: Oxford University Press;
- Lanfredini, R. (2003). Schlick and Husserl on the essence of knowledge. In P.Parrini, W.C. Salmon, M. H. Salmon, *Logical Empiricism. Historical and Contemporary Perspectives*. Pittsburgh: University of Pittsburgh Press, 43-56;
- Lanfredini, R. (2017). Anti-psychologism and neutrality: the radical empiricism of Husserl and James. *European Journal of Pragmatism and American Philosophy*, 1, 1-22;
- McCulloch, G. (2003). *The Life of the Mind. An Essay on Phenomenological Externalism*. London: Routledge;
- Millikan, R.G. (2009). Biosemantics. *Journal of Philosophy*, 86, 281-297;
- Papineau, D. (1987). *Reality and Representation*. Oxford: Blackwell;
- Schlick, M. (1938). Form and Content. An Introduction to Philosophical Thinking. In *Gesammelte Aufsätze 1926-1936*. Wien: Gerold;
- Schlick, M. (1974). *Allgemeine Erkenntnislehre*. Berlin :Springer-Verlag, 1918. Trans. *General Theory of Knowledge*. Wien-New York: Springer;
- Searle, J. (2012). *An Essay in the Philosophy of Mind*. Cambridge: Cambridge University Press;

- Husserl, E. (1966). *Die Idee der Phänomenologie*. Den Haag: Nijhoff, 1950. Trans. *The Idea of Phenomenology*. Den Haag: Nijhoff;
- Husserl, E. (1983). *Ideen zu einer reinen Phänomenologie und einer phänomenologischen Philosophie: Allgemeine Einführung in die reine Phänomenologie*, Husserliana, III/1 e III/2. Den Haag: Nijhoff, 1913. Trans. *Ideas Pertaining a pure Phenomenology and to a Phenomenological Philosophy*, First Book, *General Introduction to a pure Phenomenology*. Den Haag: Nijhoff;
- Husserl, E. (2001). *Logische Untersuchungen*, Max Niemeyer, Halle, 1900-1. Trans. *Logical Investigations*. London-New York: Routledge;
- Zipoli Caiani, S. (2019). *The puzzle of experiential justification: From qualitative states to propositional contents*. *philing* VII, pp. 127-144.



---

ROBERTA DE MONTICELLI

Università Vita-Salute San Raffaele  
demonticelli.roberta@univr.it

---

# THE PARADOX OF AXIOLOGY. A PHENOMENOLOGICAL APPROACH TO VALUE THEORY

---

## abstract

*Are values more than measures of our needs and desires or internalized social and cultural rules of behaviour, originating in cultures and devoid of any universally accessible objectivity? Is there a place for values in a world of facts? If so, how can values preserve their ideality and normativity? If not, how can value judgements be true or false? Max Scheler's Material Axiology is the best answer Classical Phenomenology provides to this dilemma. Yet Material Axiology, in particular material ethics of values, is largely ignored or looked down upon for being based on unclear presuppositions. This paper tries to provide a fresh start by clarifying the bottom-up approach characteristic of phenomenology with an exercise in experimental phenomenology in which I will analyse the actual experience of certain aesthetic values in emotionally qualified perception.*

---

## keywords

*experimental phenomenology, value theory, metaethics, material axiology*

### 1. The Paradox of Axiology

What is the Paradox of Axiology?

There is a puzzle about the very nature of values that makes the possibility of their existence seem paradoxical. It rests on an apparent *opposition between values and reality*, an opposition or contrast which philosophy has always registered. For example, it is at the root of Plato's characterization of Goodness as "beyond substance", as well as of Hume's dichotomy of facts and values and of the Naturalistic Fallacy that Moore cautions against. Even Kant's purely deontological foundation of ethics is best understood as a response to the Paradox of Axiology. The Paradox of Axiology has been rephrased, in contemporary metaethics, as the Dilemma of Metaethics, that is a dilemma concerning the status of value properties. If they were real or natural properties, they would seem to lose their essentially normative character. But, if that's so, then in order to preserve their normative character they would have to be creatures of another world. The presently available metaethical accounts of value purport to offer solutions to this dilemma.

The Paradox of Axiology is a serious one, albeit a "paradox" only in the weak sense of being something "beyond common opinion". Yet, it teaches us a lot about the very nature of values. Moreover, it calls our attention to something that is in fact part of the "whatness" of values *as they are experienced*. The Paradox existed long before its technical metaethical conceptualization was introduced. It is constantly "lived" as an tension between the ideal and the real. *Ideality is an essential feature of values*. It is also a paradoxical one, for positive values are never as vividly given as they are in the painful recognition of their absence on earth, when the corresponding goods are missing, which most of the time (but not necessarily) means that the corresponding negative values are realized in their place. Through real injustice we come to see what justice is, and that what a just society is, is what most actual societies are *not*. The *content* of justice is exemplified by what societies *could* be and *ought* to be like, or by *ideal* societies.

One of the most popular strategies for explaining this Paradox away is to remove from value terms their value *content*, or *matter*. I call the result of this move the No-Matter-of-Value Thesis. We shall explore it further below.

I shall argue for following claims:

- C1. The No-Matter-of-Value Thesis makes the Paradox of Axiology unsolvable.
- C2. Experimental Phenomenology<sup>1</sup> shows that the No-Matter-of-Value Thesis is false.
- C3. Experimental Phenomenology provides a foundation to Material Axiology<sup>2</sup>.

Most past and present conceptualizations of the Paradox are astoundingly silent about value experience and the life world, that is, about what appears to be the source of ordinary talk on value. So, in most cases, the *data* needed for any theory of value aiming to solve the Paradox are simply ignored.

In the contemporary debate on metaethics the problem is structural. For it depends on a methodological feature apparently inherited from the meta-linguistic origins of metaethics, namely, the exclusive focus on terms and concepts, as opposed to the corresponding non-conceptual contents or data. The second-order level of metaethical discussions induces authors to address the “nature” or the “status” of value properties in general, most of the time, without any intuitive exhibition and analysis of their instances. I’ll call this way of arguing at a conceptual level only, without regard for phenomenal contents, a *top-down strategy*.

Much gets lost, though, concerning the very nature of values when the fact that they are qualities that can be experienced independently of their explicit conceptualization is ignored. Take *ferocity*. Infants may be frightened by the ferocious look of a warrior’s mask, yet without having at their disposal the concepts of ferocity and of mask. This says something about the nature of values.

*Ferocity* is a typically “thick” value term. Now, the whole metaethical debate hinges on finding a satisfactory solution to the fact-value dichotomy, whether it is understood in Humean or Moorean terms. But this dichotomy, as is well known, faces apparent counterexamples in so called “thick” value-terms, or value concepts.

Such terms as “ferocious”, “courageous”, and “cruel” – or, for that matter, “graceful” or “vigorous” – denote “thick” concepts, or concepts provided with *descriptive content*, as opposed to “thin” terms such as “right” or “good” which only seem to express normativity apart from any descriptive content.

Thick concepts would appear to escape the dichotomy. Edmund Husserl discussed this issue at the very beginning of the last century<sup>3</sup>, long before it became a matter of controversy in the Fifties between Philippa Foot and Iris Murdoch on one side, and Richard Hare on the other<sup>4</sup>.

---

1 Albertazzi (2013), Husserl (1970a), pp. 203-209.

2 Scheler (1973).

3 Husserl (1970), Vol I, §§14-16; Hans Kelsen dismissed Husserl’s rejection of the dichotomy – much too hastily, I think – in his *Allgemeine Theorie der Normen* (1979), cf. Kelsen (1996).

4 Philippa Foot argues against the arbitrariness attached to the usage of “prescriptive” or “action guiding” words such as the names of virtues, or of words like “danger” or “pride” in case there where there is no “internal relation” between “commending” or evaluating and the non-evaluative meaning of these words (Foot 1978, pp. 83-104, the text was originally a paper delivered at Bedford University in 1958). Iris Murdoch opens up a much wider horizon in her criticism of “voluntarism” in value theory, targeting both (Sartrean) existential philosophy, with its reject of any rational justification of choices, as well as “linguistic” or Oxford moral philosophy, with its prescriptivism rejecting experience, research into the moral meaning of facts, learning, and discouraging of what Murdoch calls *attention*, which she explains thus: “I have used the word ‘attention’, which I borrow from Simone Weil, to express the idea of a just and loving gaze directed upon an individual reality” (Murdoch 1970 p. 33). As for Richard Hare, while his book on *The Language of Morals* (1952) reinforces Hume’s dichotomy as a distinction between descriptive and prescriptive language (the language of morals being the latter), the challenge of thick value terms is addressed in his (1981) *Moral*

## 2. Claims

## 3. The No-Matter-of-Value Thesis



Thick value terms can be used to challenge Hume's claim that one cannot derive an *Ought* from an *Is*. If, for example, being courageous is a professional quality of a warrior, then a warrior *ought* to be courageous. Or, put in terms of truth makers, there is a quality rich in descriptive content – courage – which makes an *Is* value statement true or false. That quality is the referent of such expressions as “Jack is courageous” (which is contingently true or false), and “An ideal warrior is courageous” (which is a necessary truth).

The standard move suggested by a top-down approach is to split a thick concept into two parts, i.e., a descriptive content and a normative operator, where the descriptive content is and has to be a purely “factual” or “natural” content. This was the move first made by Kelsen (in response to Husserl) and by Richard Hare (in response to Philippa Foot).

Bernard Williams approvingly sums up this move in a very clear way:

The clearest account, as so often, is given by Hare: a term of this kind involves a descriptive complex to which a prescription has been attached, expressive of the values of the individual or of the society. ... It is essential to this account that the specific or ‘thick’ character of these terms is given in the descriptive element. The value part is expressed, under analysis, by the all-purpose prescriptive term *ought*<sup>5</sup>.

On this account, value concepts have no axiological *contents*. Either they are thin, and just prescriptions, or they are thick, and all the content they have is factual. Let's call this statement the No-Matter-of-Value Thesis.

This thesis is, in fact, a consequence of Moore's thesis of irreducibility, *plus* Moore's claim that value concepts are non-analyzable. No matter how a thick value concept is analyzed in factual terms, there will be a residual value – such as the goodness of being courageous – that resists analysis in non-axiological terms. Otherwise, you run afoul of the naturalistic fallacy.

Now, this thesis is false. The rest of this paper is devoted to arguing against this very popular mistake, which I take to be caused by the kind of value blindness that goes along with a top-down approach.

#### **4. The Falsity of the No-Matter-of-Value Thesis**

Suppose you can analyse a thick value term in the way just suggested. So, a ferocious expression expresses preparedness to behave in a ferocious way, and a ferocious behaviour is such and such a behaviour, e.g., a fierce, wild, violent one. This, of course, sets us on the path of an infinite regress, for here ferocity is analysed in terms of other value qualities. Suppose, then, that ferocity just “supervenes” on a class of animal behaviours which can be described in purely “factual”, say, “natural” terms, or that the predicate “is ferocious” can be analysed in terms of those animal behaviours plus an “all-purpose prescriptive (negative) term *ought*”. Now, is there any *factual* or *natural* property shared by these animal behaviours and, say, a ferocious joke, a ferocious ideology, a ferocious question? Is there any *natural* property common to elegant things, such as an elegant dress, an elegant apartment, an elegant gait, or a piece of elegant prose? It does not seem that we can find one.

And yet, we are usually able to tell ferocious or elegant things from those which are not. How is that possible? An obvious answer is that we do so by employing the concept of ferocity. But if the preceding argument is valid, then this concept is not analyzable in terms of factual contents plus a prescriptive operator. So, the real question is: What is the source of this concept? Given that it cannot be a given matter of fact, should we not then take it to be a given

---

*Thinking* with the two-components theory (see below and footnote 5).

<sup>5</sup> Williams (1985), p. 130. This is the so-called Two Components Theory. (Putnam 2004, p. 41).

matter of value, as it were? Perhaps its source is, in other words, a non-conceptual or pre-conceptual *axiological* content or datum. Should we not then consider, at least as a hypothesis, that what a thick concept may more or less aptly capture is the *content of a given value as such*, its ideal and normative “matter”, which we have, so to speak, in front of us and “consult”?

Now, this kind of data or given non-conceptual contents is exactly what a top-down approach to value is blind to.

Our first move, then, is to make explicit the methodological rule of phenomenology as a bottom-up approach (in value theory but not in that alone):

(Meth): No theoretical problem about a type of thing *S* should be addressed without recourse to the intuitive presence of some token or instance, *s*, of *S*.

(Meth) is just an application of phenomenology’s primary charge, “Back to the things themselves”, or *the principle of the priority of the given over conceptual construction*. (Meth) is the means by which the oft-neglected world of everyday experience becomes again the privileged object of philosophical inquiry.

Our bottom-up approach can be introduced with a simpler type of case, the sort in which perception of value is in some sense “added to” perception in the most literal sense, i.e., sensory perception. We shall start with some examples of a type of quality that tradition recognizes as being “given to perception”, in *aesthesis*, that is, aesthetic qualities.

Claim 3 needs the support of experimental argumentation. To that end, some images will be shown that exemplify *ferocity* as a non-conceptual content or a *value quality*. Experimental Phenomenology will help us analyze this quality in terms of a typically tertiary quality, or *quality of demand*. This quality of demand is a global feature of perceptual configurations, constraining possible (co)variations of their contents, in all possible worlds in which ferocity is instantiated. Such a global quality *is there*, given in experience, as a *matter of value*. Its normative power, far from being “lost in description” – as if in principle *descriptive* language could not convey *ideality* – pertains to the qualified object essentially, or in all possible worlds in which it exists.

Material axiology is a generalization of this discovery, according to which value terms do have a descriptive axiological content, and a very rich one, which can be analysed by reference to objects’ axiological qualities.

Here, in Figure 1, is an example readily available on the internet of a ferocious expression. Many things are apparently ferocious, and yet it is quite difficult to find a “natural” or “factual” property that might be shared by everything that has a ferocious appearance, ranging from a warrior’s mask, certain animal behaviours, or a scene of Artaud’s cruelty theatre, to an ideology, or even some jokes. But we rightly distinguish beautiful from ugly things, cruelty from mildness. The mask depicted in Figure 1 could not possibly be deemed mild or elegant.

How can we identify cruelty and ferocity? By the concept *cruel*? But there is evidently no *factual* property shared by cruel jokes and, say, cruel meals!

Yet, as Iris Murdoch argues at length, it is always possible to improve one’s understanding of thick value concepts, such as *impudence* or *courage*. How is that possible, given what I’ve just said?

If we try to resolve the problem by appealing to the mask’s ferocious appearance, a natural answer would be that there is in its appearance a matter to be looked at – and *felt*. A matter

## 5. Back to the things themselves. The bottom-up approach of Experimental Phenomenology

## 6. Some images

of value, though, rather than a matter of fact. I've chosen a mask deliberately, for a mask is a means of make believe. When wearing such a mask, one can appear ferocious without having to be ferocious – regardless of the purpose this disguise may serve, whether it is ritual, theatrical, or a matter of carnival farce. A mask is an object under epochè, like any aesthetical object. It presents us with a quality which might be instanced in reality – the ferocious expression of an actual warrior – but which we grasp, so to speak, in the “abstract”, divorced from any genuine exercise of ferocity.



Figure 1

Let's briefly practice “contemplation,” or, in Murdoch's terms, the improvement of our *insightful* understanding of ferociousness. We certainly have a lot of matter to conceptualize, matter that is first given to visual-cum-emotional perception. We are “struck” by more and more features, which we may or may not be able to adequately capture in words (not just *any* word will do!). We detect a quality of *wildness* or *savageness* which suggests, despite the evident shape and aspect of a human face, a *lack of humanity*. We sense *hostility*, *aggressiveness*, *rapaciousness* and *greed*. We are scared by the face's *fury* (seen in the eyes) and apparent *mordacity* (visible in the teeth). The grim look of the mask exhibits a power of *tearing apart* and destroying in its sharp and pointed features...

This description serves as a counterexample to the idea that value terms are just action-guiding terms. When I describe this mask as “ferocious” I am not warning you against any menace to your life or integrity, nor am I prescribing you to avoid encounters with it, etc. The point also works as a counterexample both to Hare's prescriptivism and to emotivism. For the very same quality might be manifested in an act of cruelty, too. You would no doubt be right to fear being killed. And, were that to actually happen, it would be awkward if we then had to choose between either (a) admitting that “It was a cruel, a ferocious act” is a value judgment while denying the judgment a truth value or (b) admitting that the judgment corresponds to the pertinent facts while denying it the status of a value judgment. That awkwardness indicates, as is usually the case with thick terms, the untenability of the fact/value distinction. So, here a Two-Component Theory à la Hare (ascribing the judgment a descriptive content plus a prescriptive force) might help. But again, it would be no help for understanding the related case of the mask, despite the fact that the value quality manifested there is exactly the same.

Expressive qualities are a subclass of value qualities that typically “present” the agents having them as other *selves*, namely, as subjects (or quasi-subjects) of emotion and action. It is by

means of these qualities that we usually take something to be another self. They present us with socially relevant properties of encountered agents and their actions. But, when experienced in a purely aesthetical attitude, expressive qualities can be explored in a “detached” way, so that the pertinent emotions can be vividly felt without becoming *reasons for action* (think of thrillers, horror pictures, and so on). Aesthetic experience switches off action and at the same time deepens cognition. Figure 2 gives a further example, this time presenting a familiar scenario from ordinary life, of the “abstraction” involved in the appreciation of aesthetic value.



Figure 2

In comparison with the previous examples, the Halloween Pumpkin is a more “abstract” version of a ferocious expression. It is an almost simplified or schematic version of ferocity (an “*eidōs*” of it?). You’ll notice nonetheless how well the salient features of apparent ferocity are preserved. In fact, a matter of real life and action has become pure play here with the help of this “aesthetic” object and our taking on the “right” attitude toward it. In aesthetic experience one apprehends the quality without being moved to action, thus taking it up in a dispassionate cognitive attitude. In pretense, as in some kinds of plays and games, the motivational power of qualities is in the “as if” mode, but you enact corresponding “as if” actions in relation to them. Aesthetic experience switches off real action and switches on value cognition. This is especially true in the case of expressive qualities. A brief foray into the history of 20<sup>th</sup> century visual art will make the point clear.

In the Preface of an old English translation of Vassily Kandinsky’s *Punkt und Linie zu Fläche* (1926) we find a nice anecdote on the meaning of “abstract” painting:

Upon his return to Munich, one evening there occurred at dusk the magical incident of his seeing merely the form and tone values in one of his paintings. While not recognizing its subject, he was not only struck by its increased beauty but also by the superfluity of the object in painting, in order to feel its spell. It took him fully two years to crystallize this miraculous discovery<sup>6</sup>.

---

<sup>6</sup> Kandinsky (1926), p. 7. The date of this discovery is supposed to be about 1908. Even if Rebay does not report this detail, the tradition has it that the picture that struck Kandinsky in this way was not one of his own, but rather Monet’s *Sheafs in the sun*. The (1926) treatise was published sixteen years after his first one, *Das Geistige in der Kunst* (1910), and is based on Kandinsky’s teaching at the Bauhaus in Berlin (1921-24). It’s a nice coincidence that while the first treatise, written in Munich, is more or less contemporary with the flourishing of the Munich psychologist Theodor Lipps’ studies on empathy and aesthetics and the critical, very productive discussion of them within the Phenomenological Circle of Munich, animated by Scheler and attended by two of the most significant phenomenologists of aesthetics, Moritz Geiger and Dietrich von Hildebrand, the second one, crystallizing his teaching

The quotation conveys all that is needed to appreciate the series of visual experiments Kandinsky provides in the appendix of this remarkably valuable work. Let's examine just two examples. So-called "abstract" painting only makes explicit – and does so in a programmatic way – what has always been true of painting, regardless of the views the painter or mainstream opinion about painting in different epochs. It highlights how visual art, correctly enjoyed, lets us see the *how*, and not only the *what*, of visible things. Generalized, we could say that aesthetic information is about the *how* and not the *what*, or that it's about *whatness*, but only insofar as it can be made apparent independently of whether and where it exists, or whether it is really as it appears. Painting is about qualities, even those making up the solidity and three-dimensionality of the painting's subject matter. It is about the gravity and seriousness of what is real, as is particularly evident in Cezanne's use of the canvas. Qualities are the real "subject" of even the most figurative and "classic" painting. Geometry itself appears as the quality of an orderly world, an intelligible cosmos encompassing the City of Humans. That is true, for instance, of certain paintings by Piero della Francesca.

Once you realize that the visual arts help to free visual perception from the practical tasks it serves in ordinary life and to free visual contents from their function of orienting you in reality, the supporting role of the "subject" (e.g., a depicted object) in such work is no longer necessary for appreciating the "how", the pure visual content of possible perceptual worlds. Indeed, there is no need for a return to Platonism to understand Klee's dictum that painting does not reproduce the visible, but makes even the "invisible" visible.

Klee's dictum refers to what I've called value cognition. By "contemplating" visual art we come to "improve our understanding" (as *per* Murdoch) of expressive qualities. Actually, we come to improve our discriminative perception of them. More examples will prove helpful here.

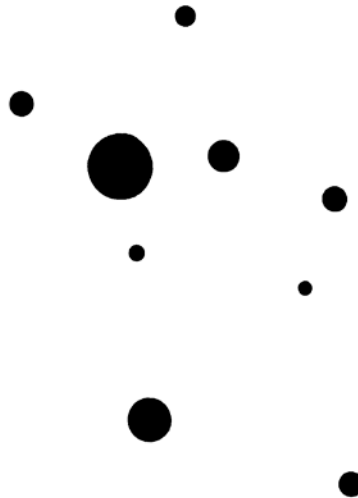


Figure 3. From V. Kandinsky, *Point and Line to Plane*<sup>7</sup>

---

years in Berlin, is more or less contemporary to the flourishing of the Berlin School of Gestalt Psychology in the 1920s and the beginning of the 1930s, before the Nazi catastrophe.

<sup>7</sup> Kandinsky (1926), *Anhang* (Appendix), Diagram 20, (1946) p. 152.



Figure 4. From V. Kandinsky, *Point and Line to Plane*<sup>8</sup>

The “titles” given by Kandinsky to these drawing are respectively: “9 points in ascent (emphasis upon the diagonal d-a through weight)” (Figure 3<sup>9</sup>); and “Diagonal tensions and counter-tensions with a point which brings an external construction to inner pulsation” (Figure 4), The first is meant to illustrate the role of the point in visual form and the latter the role of the line.

Both configurations feature a striking *dynamism*, a quality we perceive or feel even if we know that the lines or the points on the paper are in fact motionless. Theodor Lipps (1906) describes this kind of quality as presenting “the life of lines”, as it were:

[A] movement, such as stretching out, growing longer, self-restricting, abruptly starting and ending, or steady sliding, swinging up and down, bending, stooping, squeezing and expanding. All these predicates do not refer to geometrical features of the form, but denote activities...<sup>10</sup>.

We can call the qualities Lipps refers to as *dynamic qualities*. The drama of the lines and the rising of the points in Kandinsky may well be a good (almost) contemporary example of what Lipps has in mind<sup>11</sup>.

<sup>8</sup> Kandinsky (1926), Anhang (Appendix), Diagram 20, (1946) p. 186.

<sup>9</sup> Kandinsky rejects customary distinctions of “the graphic” and “the pictorial”, according to which line is the essential element of graphics but is “in painting, contrary to its nature and, therefore, forbidden”. Here is his rationale: “This is a characteristic example of the existing confusion in concepts: that which can easily be segregated and placed in separate categories is mixed together (art, nature), and, on the other hand, the things that belong together (painting and graphics) are carefully separated from each other. The line is considered here to be a ‘graphic’ element and not to be used for ‘pictorial’ purposes, although an elementary difference between ‘graphics’ and ‘painting’ cannot be found and could never be established by the theorists mentioned”. Kandinsky (1946), p. 110.

<sup>10</sup> Lipps, T. (1903-06), p. 184. Our translation.

<sup>11</sup> Even if the theories of *Point and Line to Plane* (1926) emerge after Lipps’ *Aesthetics*, Kandinsky’s “first abstract watercolour” was painted about 1908, preceded by a rich series of works where the dedicated study of both expressive and dynamic qualities is more than evident. The same is true of other work done within the *avant-garde* art of the early 20<sup>th</sup> century, and especially in Munich, where Kandinsky, with Franz Marc, Paul Jawlensky and others founded, as is well known, the *Blaue Reiter* movement (1911-14).

**7. Abstract  
Visual Art and  
Experimental  
Phenomenology**

Instituting a pictorial space means literally “abstracting” these pervasive yet silent qualities from our concrete surrounding world in order to display them before our eyes, as objects whose variations and possibilities offer us a new, infinite domain of exploration<sup>12</sup>. This is, after all, what painters have always done. The frame of a painting manifests the initial “bracketing” the painter performs with respect to all the non-aesthetic goals of ordinary perception. It separates pictorial space from one’s surrounding actual space.

Yet it is only around the birth of “modern” art that this “bracketing” and this “abstraction”, achieved with the aid of aesthetic objects and their means (scores and melodies<sup>13</sup>, paper and drawings, collages, etc.), entered the laboratories of science and gave rise to experimental phenomenology<sup>14</sup>, including, very early on, Gestalt psychology.

What grounds abstract painting is indeed the very same discovery with which Gestalt psychology began. Both recognize that perceptual contents, far from being the unorganized “multiplicity” or “chaos” (Kant) of sense data postulated by the empiricist tradition, are organized by “configural” or structural properties. These properties are given, not “constructed”, and are non-conceptual, pre-linguistic, and often multi-modal or amodal in nature.

The fact that the two groups of researchers share a common root is something no author could better bear witness to than Rudolf Arnheim, the brightest pupil of Max Wertheimer. Arnheim was a brilliant young art critic in the roaring 1920s in Berlin, the director of the Italian Istituto del Cinema in the 1930s, and the founder of the psychology of visual art in his post-exile American academic life<sup>15</sup>. I quote him at length below not only because he and others of his school gave us a language for describing the axiological contents of aesthetical values, but also because there are passages in his work that bring the common root of Gestalt Psychology and abstract painting to the foreground. Take, for instance, this powerful synthesis of the “Kandinskian” analysis of dynamic qualities we’ve just reconstructed:

Visual experience is dynamic. This theme will recur throughout the present book. What a person or animal perceives is not only an arrangement of objects, of colors and shapes, of movements and sizes. It is, perhaps first of all, an interplay of directed tensions. These tensions are not something the observer adds, for reasons of his own, to static images. Rather, these tensions are as inherent in any percept as size, shape, location, or color. Because they have magnitude and direction, these tensions can be described as psychological ‘forces.’<sup>16</sup>

---

<sup>12</sup> This point of course generalizes across the visual arts, and, *mutatis mutandis*, the institution of other aesthetical spaces, like that of music.

<sup>13</sup> As is well known, the notion of “Gestalt” had been introduced into psychology by Christian von Ehrenfels in 1890 in his essay “On Gestalt qualities”, where, observing that humans can recognize two melodies as identical even when no two corresponding notes in them have the same frequency, he argued that these forms must possess a “Gestalt quality”—a characteristic that is immediately given, over and above the single tones. Cf. von Ehrenfels C., *Über Gestaltqualitäten*. (1890), *Vierteljahrsschrift für wissenschaftliche Philosophie*. 1890;14:224–292. (Translated as “On Gestalt qualities”). In B. Smith (Ed. & Trans.), (1988), *Foundations of Gestalt Theory* (pp. 82–117). Munich, Germany/ Vienna, Austria: Philosophia Verlag.)

<sup>14</sup> I use this expression, “experimental phenomenology,” to refer to that part of early experimental psychology that harboured and outlasted classical Gestalt psychology, on the one hand, and that instigated the powerful approach to generalization which led Husserl to the study of the eidetic universe. See De Monticelli (2018).

<sup>15</sup> Rudolph Arnheim’s life (1904, Berlin – 2007, Ann Arbor, USA) actually spans the great temporal distance between the time of Stumpf, Husserl and Wertheimer and our own. See Arnheim (1954, 1974). The 1974 edition is a revised version of the original, published after 20 years of teaching in several New York Universities and Harvard University.

<sup>16</sup> *ibid.*, p. 11.

Perhaps no image could better illustrate the concept of visual force or tension at the heart of Kandinsky's theory of elements (especially as it concerns point and plane) than the one reproduced in Figure 5. It exemplifies one of the phenomena to which Arnheim devoted most of his research, namely, the "power of centre."

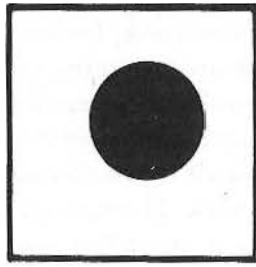


Figure 5. The Power of the centre

This figure – which is the first presented in his main work, occurring in the section titled "Balance"<sup>17</sup> – illustrates the first principle of Gestalt theory. The fact that we perceive the slightly off-centre position of the "point" on the square plane is evidence for the principle that we grasp organized wholes. What we experience in perception is not the association of atomic perceptual data, but the immediate givenness of the elements appearing from the start as a function of the wholes to which they belong.

Wagemans and colleagues put the point thus: "The contents of our awareness are by and large not additive but possess a characteristic coherence"<sup>18</sup>. There are several additional principles further articulating this notion of coherence that are also illustrated by this simple figure. Firstly, there is the principle of figure/ground, according to which the partial contents we perceive are experienced as segregated from a background. Secondly, there is the principle of unity, the basic organizing factor discussed by Kandinsky, which highlights structural features such as the square basic plane with its horizontal and vertical boundaries, its center, and its diagonals. Here the invisible center is manifested by the off-center point, and induces a dynamic quality within the figure, giving the impression of things being imbalanced.

I defer again to the expertise of Arnheim:

The disk in Figure 5 is not simply displaced with regard to the center of the square. There is something restless about it. It looks as though it had been at the center and wished to return, or as though it wants to move away even farther. And the disk's relations to the edges of the square are a similar play of attraction and repulsion<sup>19</sup>.

The "power of the center" in a sense "requires" the disk to be centred, that is, for the simple composition to reach *balance*. This is the first mention I've made so far of that feature of the perceived world that Wolfgang Koehler called "requiredness", a type of "oughtness" which turns out to pervade the life world and also indicates for us the "place" that values have in it. This simple example shows that it is not really a "subjective" – or rather, arbitrary – matter

---

<sup>17</sup> Ibid. p. 10.

<sup>18</sup> Wagemans and colleagues, 2012: 1172–1217. Published online 2012 Jul 30. doi: 10.1037/a0029333, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3482144/>

<sup>19</sup> Arnheim (1974), p. 12.



to perceive what is required in a given context. Requiredness is not reducible to the allegedly prescriptive component of a thick value term (like *imbalanced* or *disharmonious*). There is of course nobody “requiring” the disk to be centred, any more than there is always someone commanding a warrior to be courageous or commanding someone giving an argument to ensure its validity. Nor are such thick value terms necessarily action guiding. They have plenty of descriptive content. But this content consists in qualities of requirednesses or – as it is referred to in experimental phenomenology – of “qualities of demand” (Albertazzi 2013). We literally perceive the demands that things manifest.

Similarly, the apparent ascending movement we perceive in Kandinsky’s rising points in Figure 3 is also not “merely subjective”. If it is an illusion, it is not correctable, much like those so called “illusions” (such as the Muller-Lyer) that once led Gestalt psychologists to radically question the causal-physicalistic model of perception. The lightness of the “ascending points” is a material value quality, the global quality of a well-organized whole.

**8. Conclusion** Moore’s greatest intuition – that ideality is irreducible to reality – is tangled up with his worst mistake, which was to hold that Goodness or the Good is “a simple, non-analysable object of thought”<sup>20</sup>. Against Moore, we ought to hold instead that “Goodness” is a proxy word for “any positive value quality”. That is, it functions as a variable ranging over thick or material values. As an attributive adjective applicable to exemplars of various kinds of things (e.g., as with the phrase “a good knife”) it is a variable ranging over the positive value qualities of ideal exemplars of the relevant kind (e.g., functional value qualities, in the case of this domestic tool). As a moral predicate (“morally good”) it is a thick, –not thin, concept, ranging over all moral virtues and qualities of an intention, action, or person enabling them to realize the best value(s) possible in a given situation, which presupposes that we are capable of perceiving what is required in the first place, in short, that we are capable of *attention*.

The upshot of this analysis is that, more generally, thin values depend on thick values, and it is false that thick values have no descriptive *value* content, i.e., that value qualities cannot be analysed, explained or described. They can, but only in terms of other values, contrary to Moore’s ineffability thesis.

Hare and Williams make the complementary mistake of reducing thick value terms to their alleged “real-descriptive” content plus a universal or indeterminate commendatory force.

The No-Matter-of-Value thesis makes material axiology inconceivable.

But I hope to have shown not only that counterexamples to that thesis are *conceivable*, but also that we can *see and feel* what we conceive in them, thus making it possible for us to describe them more or less adequately and supply Material Axiology with its grounding evidence.

#### REFERENCES

- Albertazzi, L. (2013). *Handbook of Experimental Phenomenology, Visual Perception of Shape, Space and Appearance*, Oxford: Wiley Blackwell;
- Arnheim, R. (1954, 1974). *Art and Visual Perception. A Psychology of the Creative Eye*. London and Los Angeles: California University Press;
- De Monticelli, R. (2018). *Il dono dei vincoli – Per leggere Husserl*, Milano: Garzanti Libr;
- Foot, P. (1978). *Virtues and Vices* Berkeley: University of California Press, Ch V, Moral Beliefs;
- Hare, R. (1981). *Moral Thinking*, New York: Oxford University Press, 1981;
- Husserl, E. (1970a). *Philosophie der Arithmetik. Mit ergänzenden Texten (1890-1901)*. Edited by Lothar Eley. The Hague, Netherlands: Martinus Nijhoff;

---

<sup>20</sup> Moore (1903, 1959), §15, p.21

- Husserl, E. (1970b) *Logical Investigations*, Engl. Transl by J.N. Findlay, Routledge, London and New York; Vol I, *Prolegomena to Pure Logics*; Vol. II, Third Investigation;
- Kandinsky, V. (1926). *Punkt und Linie zu Fläche*, Engl. Transl. (1946) *Point and Line to Plane. Contribution to the Analysis of the Pictorial Elements*, edited and prefaced by Hilla Rebay. New York: Solomon R. Guggenheim Foundation;
- Kelsen, H. (1996). *Théorie générale des normes*. Paris : PUF, Ch. 52. « Le ‘contenu théorique’ de la norme selon Husserl », pp. 269-272 (Trad. Fr. of Kelsen (1979), *Allgemeine Theorie der Normen*, Wien : Manz Verlag);
- Lipps, T. (1903-06). *Ästhetik. Psychologie des Schönen und der Kunst*. Hamburg/Leipzig: Voss. Partial tr. It. *Estetica*, ed. by A. Pinotti (1997). Milano: Guerini;
- Moore, G.E. (1903). *Principia Ethica*. Cambridge: Cambridge University Press 1959;
- Murdoch, I. (1970). *The Sovereignty of Good*. London and New York: Routledge;
- Putnam, H. (2004). *Fatto/Valore. Fine di una dicotomia*. Roma: Fazi, It. tr. of (2002), *The Collapse of the Fact/Value Dichotomy and Other Essays*;
- Scheler, M. (1973). *Formalism in Ethics and Non-Formal Ethics of Values*, Engl. Transl. by M Frings and R.L. Funk. Chicago: North Western University Press;
- von Ehrenfels, C. (1890). Über “Gestaltqualitäten”. (1890), *Vierteljahrsschrift für wissenschaftliche Philosophie*, 14: 224-292. Transl. “On ‘Gestalt qualities’”. In B. Smith (Ed. & Trans.), (1988), *Foundations of Gestalt Theory* (pp. 82-117). Munich, Germany/Vienna, Austria: Philosophia Verlag;
- Wagemans, J. and colleagues, (2012). A Century of Gestalt Psychology in Visual Perception I. Perceptual Grouping and Figure-Ground Organization. *Psychological Bulletin* 138(6): 1172-1217. doi: 10.1037/a0029333, <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3482144>;
- Williams, B. (1985). *Ethics and the Limits of Philosophy*. London and New York: Routledge.



# SECTION

# 4

# SECTION 4

PHILOSOPHY AND GENDER ISSUES

*Marina Sbisà*

Ways to Be Concerned with Gender in Philosophy

*Laura Caponetto*

Filosofia del linguaggio femminista, atti linguistici e riduzione al silenzio

---

MARINA SBISÀ  
University of Trieste  
sbisama@units.it

---

# WAYS TO BE CONCERNED WITH GENDER IN PHILOSOPHY\*

---

## *abstract*

*In this paper, some (non-exhaustive) ways in which philosophy can tackle gender are presented and examined. Starting from mainly negative critical stances and proceeding towards more positive approaches, the following ways to tackle gender are distinguished: the critique of gender through discourse analysis; the critique of gender essentialism; research on how gender concepts work; reflection upon how gender issues relate to issues of intersubjective recognition. The first three ways, although giving important contributions about how gender is conceived of, do not seem to give firm grounds to the rejection of gender discrimination, the refutation of essentialist beliefs about the genders, and the neutralization of the normative import of gender concepts. As a fourth way, I propose considering gendered subjectivity in the context of the dynamics of intersubjective relations, framing it in a view of intersubjective recognition as a basic process in human life. In this perspective, gender issues constitute a context in which people's capability and willingness to recognize others as subjects, and specifically human persons, is tested.*

---

## *keywords*

*critique of gender stereotypes, gender essentialism, gender discrimination, recognition of the other's subjectivity*

\*I thank Claudia Bianchi for urging me to write this article, an anonymous referee for his or her comments and suggestions, and Laura Caponetto for insightful discussions.

---

**Introduction** In this paper, I illustrate some (non-exhaustive) ways in which philosophy can tackle gender. My aim is not to provide an overview of philosophical work on gender and gender-related issues in the past decades,<sup>1</sup> but to single out some kinds of approach to gender, whether actual or possible, and examine the difficulties or limitations of each. Starting from mainly negative critical stances and proceeding towards more positive approaches, I distinguish:

1. the critique of gender through discourse analysis;
2. the critique of gender essentialism;
3. research on how gender concepts work;
4. reflection upon how gender issues relate to issues of intersubjective recognition.

Each of the four main sections of the paper is devoted to outlining one of these kinds of approach. Since there are close connections between the first and the second, and between the second and the third, these three can be considered as following up or elaborating on the preceding one. The third way to tackle gender draws on recent research concerning the linguistic expression of categorizations, to which gender categories belong. This research, although not specifically about gender, offers a way in to understanding the normativity of gender models as well as gender-based discrimination. As to the fourth way, I propose considering gendered subjectivity in the context of the dynamics of intersubjective relations, framing it in a view of intersubjective recognition as a basic process in human life. Hopefully, this kind of approach might shed some new light on the debates on gender and gender-related issues.

**1. Critique of gender through discourse analysis** Because of the complexity of gender-related issues in contemporary society, and the fact that assumptions and expectations about gender are usually tacit, it is not easy to spell out what gender is, or even simply what gender (and genders) are commonly held to be. One instrument at hand for obtaining some clarity about how we think about the matter is the examination of how gender is dealt with in discourse and other social practices. Although one may object

---

<sup>1</sup> So, I will not comment upon feminist philosophical work done in the various areas of philosophy (such as that presented and discussed in Fricker and Hornsby, 2000, and Saul, 2003). I believe, however, that distinguishing between kinds of approach to gender as I will be trying to do is indirectly relevant also to that kind of philosophical work.

that such a task does not pertain to philosophy, I maintain that it does (in part at least), for two reasons. Firstly, let us not forget that at a certain time in the 20<sup>th</sup> century philosophy was described as an activity of clarification (under the influence of Wittgenstein: see e.g. 1922: §4.112; 1953: §133) and even though this characterization is now taken to be outdated, I believe that it might still be worth consideration as non-exhaustive, alongside more ambitious ones (for which see e.g. Williamson 2006, 2007). The idea of philosophy as an activity of clarification is traditionally connected to an antimetaphysical attitude, so that what is expected to be clarified is philosophical discourse itself, insofar as it expresses unsolvable “philosophical perplexities”. In my opinion, however, the idea can quite sensibly be extended to include all those issues we find difficult to tackle because of underlying conceptual confusion or lack of explicitness. In this sense, though seemingly a matter of applied linguistics, or somewhere between linguistics and sociology, discourse analysis (see Fairclough 1995, Wodak and Chilton eds 2005) can also be practised in a philosophical spirit, not only because it sometimes profits from the use of instruments coming from the philosophy of language such as speech acts, presupposition and implicature, but because it throws light upon the contents and organization of our ideas.<sup>2</sup> Secondly, once our tacit assumptions about some complex and controversial issue and their implications are made explicit, the question remains open as to whether such assumptions are acceptable or even correct. It seems to me that the critical assessment of commonly held assumptions about gender and genders is to a large extent a philosophical task, connected to the understanding that human beings have of the human condition and one to be pursued by argument. Even if it goes beyond mere clarification, this task is at least contiguous to clarification activities, because it is made possible by the explicitness that follows from them. We cannot reject an assumption about gender (or criticize it or argue against it) unless we are aware of it, and we might not even realize it needs to be criticized unless we are aware of the precise role it plays in our discursive and social practices.

Many of the assumptions about gender revealed by analyzing discourse and observing social practices have the controversial feature of being “essentialist”, that is, they present themselves as concerning “essential” properties of a certain gender-related class of human beings. Now on the one hand, were this true, alleged members of the class who do not possess such properties should not be counted as real members; on the other, accepting people who lack those properties as members of the class amounts to falsifying the claim that the properties are essential. But this is not what actually happens. Essentialist assumptions bear *de facto* normative implications, that is, the members of the class of human beings at issue, provided they want to be (or cannot avoid being) members of that class, are accepted as members, expected to have or develop or acquire the alleged essential properties, and negatively evaluated as not “normal” or not “good” members if they do not. Claims about the alleged “essence” of a gender thereby become models to which people who are assigned that gender have to conform. For example, it is often held to be part of the essence of a gender that people belonging to it have certain abilities and typical dispositions, which make them suited to certain roles in society, and this can create in them a normative orientation towards actually assuming those roles.

## **2. Critique of gender essentialism**

---

<sup>2</sup> Aspects of the use of language that are revealing of gender models are studied in the pioneering work of Robin Tolmach Lakoff *Language and Woman's Place* (1975). For a use of speech act theory and of presupposition applied to gender-related issues, see for example Hornsby (1994); Hornsby and Langton (1998); Langton and West (1999). I have used the instruments of pragmatics to make explicit various assumptions about or attitudes towards the feminine gender in my research on Italian women magazines (1976) and on books on pregnancy and birth addressed at women (1984).



A further essentialist assumption is that belonging to a gender is a fact determined by nature or even a fact of human nature (and thus part of the essence of human beings). These latter assumptions add to the normativity of gender models the compulsion to follow one of them: if belonging to a gender is an essential part of being human, then nobody can escape gender altogether, which means that every one of us will be subject to one or other of the gender models available.

Assumptions about gender and genders can be criticized as empirically false generalizations, and they can sometimes be proven to be such. But irrespective of their empirical truth or falsity (namely, whether the majority of the members of a gender actually display the properties supposed to be “essential” to it), they can (and should) be criticized as essentialist, which involves rejecting the idea that they describe the nature or essence of a gender, thereby limiting or suspending their normative import.

In their attempt to look at women’s difference afresh, and therefore not merely as the lack of allegedly masculine virtues but as rich in positive contents, some feminist thinkers have raised claims about the feminine gender that appeared to have essentialist implications (Irigaray 1977; Gilligan 1982).<sup>3</sup> Many other feminist thinkers, on the other hand, as well as many researchers in linguistics and sociology, have chosen to criticize and reject gender essentialism. Indeed, gender models are quite obviously cultural constructions<sup>4</sup> and the rhetoric of naturalness is, therefore, both false and politically detrimental. But this anti-essentialist stance is not without problems.

To cope with the fact that most human beings have unambiguously sexed sets of chromosomes and bodies, anti-essentialists have used two strategies, one of which involves introducing the sex/gender distinction (see Lorber and Farrell eds 1991), while the other radicalizes anti-essentialism, considering sexes too as cultural constructions (see, most famously, Butler 1999). As we know, the sex/gender distinction assumes that in the opposition between male and female, what is natural can be distinguished from what is cultural (and therefore, in principle, subject to choice). Unfortunately, though, it is empirically difficult, not to say impossible, to tell the difference between what people do because of their nature and what they do because of education, habit, tradition, and the like. Indeed, what appears to be cultural may be analyzed in a “naturalized” way, while what appears to be natural can almost always be modified by training or, when appropriate, by the use of technology. The sex/gender distinction, then, although at first sight clear and almost commonsense, actually turns out to be weak, and this weakness leads anti-essentialists to favour radical solutions, claiming that sex is as cultural as gender after all and that the sex/gender distinction does not hold. That all difference between human males and females is a matter of performance, however, thus belonging to the realm of culture, is not thoroughly convincing. One may wonder, for example, whether a woman’s failure to display certain properties belonging to the feminine gender (for example, a girl ignoring dolls and playing with cars and trains instead) and a woman undergoing surgery to change her sex are to be reasonably considered as graded manifestations of the same kind of process. The toughest point in the debate about essentialism is, however, reproduction. What about reproductive roles in human beings? Firstly, what about the physiological roles that male and female play in reproduction? From

---

<sup>3</sup> *Speculum* (Irigaray 1974) is not an essentialist book, but in later works (among which Irigaray 1977) this author seems to derive the properties of the feminine subject directly from features of female sex. However, see Braidotti’s comments (1991: 248-263). Although Gilligan is often treated as the prototype of essentialist feminism (see for example Antony 2012), the interpretation of her work is controversial (see Saul 2003: 216-220).

<sup>4</sup> This assumption works quite well as a framework for the investigation of gender identities and relations in social and communicative interaction. See for example Kotthoff and Wodak (eds) (1997), Baron and Kotthoff (eds) (2001).

the biological point of view, human reproduction resembles that of other mammals: male and female reproductive roles are not interchangeable nor symmetric, and (for the time being at least) the functions played in the reproductive processes by male and female individuals can only be partly replaced by technological aids. Secondly, does evolutionary adaptation to male and female reproductive roles determine other properties of male and female individuals? Perhaps not at all, or perhaps to some extent only.

Granted, technological progress can turn biological functions into cultural choices. This is apparent nowadays in human reproduction, particularly with respect to motherhood: thanks to technology, women now have the means to decide whether the sexual intercourse they might want to have will be potentially fertile or not, so that conception (and therefore motherhood) can no longer be considered as matters of biological fatality. Other more complex technological interventions make it possible to replace intercourse with artificial insemination, or let a woman host an embryo of which she is not the genetic mother; moreover, after a certain minimum gestation time, foetuses can now survive and develop outside the mother's womb. All these new techniques open the way to choices that were not "naturally" available, such as having a child from a dead husband, or getting pregnant after the end of one's childbearing age. In these cases, the obvious contrast between what is done by means of technology and what would happen according to "nature" suggests that on each occasion we face a choice between a technology-guided and a nature-guided line of conduct. But these two kinds of behavior are not "chosen" by people in the same way and to the same extent. In most cases, it is reasonable to assume that what happens without technological intervention has a default status: the issue of choice only arises when something (a precise goal, for example) makes the use of technology pertinent, and so the two situations are not symmetrical. So, while whatever technology does rather obviously pertains to culture, it does not follow that what happens without the use of technology should be counted as cultural as well. After all, it seems only fair to concede that the body enjoys a sort of givenness which philosophy should not sweep under the carpet but seriously thematize.<sup>5</sup>

It is quite understandable that people seeking to free themselves from gender models should attempt to shift the biology of reproduction from nature to culture: taking it all to be a cultural matter removes from gender models any possible biological legitimization. There might be other ways of rejecting gender essentialism, however, which do not need such a move. For example, it should be possible to see the nature/culture distinction as a continuous line between two poles, and recognize, in contrast to the way the distinction is construed when turned into a dichotomy, that culture has its determinism and nature its flexibility. As to our bodies, it should be possible to take them as starting points (be these resources or constraints), namely as things that do not speak for themselves but which we elaborate on and assign significance to. Of course, it is no easy task to construct a defense from essentialism along these lines, especially due to the widespread expectation that any theory that is not anti-essentialist must be essentialist, and *vice versa*.

Given the difficulties in dealing with essentialist assumptions about gender, one would expect that looking at language again would be useful, no longer to specify assumptions about gender making them available to criticism, but to investigate how it is that these assumptions are so pervasive of our experience and activity and tend to be perceived as normative, making one's

### **3. How do gender concepts work?**

---

<sup>5</sup> More attention to this theme has been paid by phenomenology: a very interesting example is Heinämaa (2014). In my own work (Sbisà 1996), I have put forward a conception of feminine subjectivity as developing in the elaboration of female sex-related experiences.

gender appear as one's inescapable fate. Another way to concern ourselves with gender issues in philosophy might therefore be to discuss, not so much the content of gender concepts, as their nature or structure, and the way they work in cognition and communication.

Let us first examine which ways of conceptualizing gender are available in principle. To begin with, gender can be (and often is) conceptualized as dichotomically exhaustive: there are two genders and each individual has to belong to one of them, or even, every property that an individual may have must be part of one or other of the two gender concepts. In this case, individuals have to belong to a gender, and if the latter assumption is also taken to hold, this gender will determine all their properties. This is the most restrictive way of conceptualizing gender. Emblematic of this is the astonishing claim I happened to come across recently on a poster promoting protest against the alleged influence of so-called "gender theory" in Italian educational programmes: "boys are male and girls are female".<sup>6</sup>

To lessen the compulsoriness and pervasiveness of gender, one may conceive of gender concepts as comprising only gender-relevant properties. This makes it possible, at least in theory, that an individual lacking those properties might fail to belong to a gender. Moreover, properties not relevant to gender could be possessed by individuals of any gender and therefore, for example, shared by individuals belonging to different genders. It seems to me that our usual way of conceiving of gender matches this model at least in this respect. Compulsoriness of gender would be also mitigated if a society and its language were to allow for an open number of genders: this would be made possible if the contents of existing gender concepts were conceived of as non-exhaustive of the possible sets of gender relevant properties. However, this way of conceiving genders is not without risks, since it fosters the individual's identification with a gender, thereby indirectly confirming that gender belonging is compulsory after all.

As a further alternative, gender belonging may be conceived of as gradable. In this perspective, one individual can be deemed more masculine than another, or even more masculine than feminine (or the other way around), depending on the number of properties he/she possesses belonging either to the masculine gender concept or to the feminine gender concept. Those properties are still either feminine or masculine, and it is their proportion that determines gender belonging. Adopting this perspective enables us to talk of more (or less) feminine women and more (or less) masculine men, and although we may be accepting actual people as "hybrid", the fixed borders and received contents of gender concepts themselves are confirmed.

This brief exploration of possible ways of conceptualizing gender is perhaps enough to show, or at least to strongly suggest, that none of these ways is quite free from the pervasiveness and compulsoriness that make gender concepts such a source of trouble in real life. Clearly, the solution to these troubles cannot be found in the structure of gender concepts considered in isolation, and so it is worth considering whether problems with gender concepts may be rooted in general facts about language and cognition.

Gender labels are usually employed to make generic statements such as "Men love soccer" or "Women are emotional". Recent research on generics, by Sarah-Jane Leslie in particular (Leslie 2007, 2008, 2017),<sup>7</sup> maintains that such statements are not equivalent to quantified statements, but are held to be true when the property assigned to the category is prevalent in the category or else when, even if not prevalent, the property is characteristic of the category or striking for some other reason (e.g., members of the category that possess it are dangerous,

---

<sup>6</sup> I saw this poster in Rome in September 2017.

<sup>7</sup> For an overview, see Leslie and Lerner, 2016.

as in “Mosquitoes carry the West Nile virus”). Elaborating upon various experimental findings (showing, for example, that children understand generics as generalizations about kinds or categories before they fully master quantified statements, and that people tend to recall universally quantified statements as generics),<sup>8</sup> Leslie claims that generics are the most basic way in which cognitive generalizations are expressed in language. So, generics support the fundamental cognitive activity of categorization both in the area of natural sciences (notably biology) and in social cognition, but at a price: the use of generics appears to regularly activate essentialist beliefs,<sup>9</sup> that is, the category is taken to be grounded in an essence, which means that members of the category are taken to share something that causally grounds their manifestation of certain properties (Leslie 2017). It should be noted that this use of generics does not suggest that the category they are about has fuzzy borders: rather, it makes the classic view of a category as having a perfectly delimited extension compatible with the possibility (and actual existence) of exceptions. Accepting a generic disposes the receiver to infer that any arbitrarily chosen member of that category is endowed with the property expressed by the predicate of the generic. It seems to me that the essentialist beliefs observed by Leslie in association with the use of bare plurals as subjects of generics could also be described as existential presuppositions (and accommodations thereof): the category is presupposed to “exist” as a category, and therefore, to have an essence, and the receiver is implicitly invited to share such a background assumption. It is insofar as a certain category is assumed to exist and be causally grounded (so that its members at least share the disposition to manifest certain properties), that it (or its members, *qua* its members) may become the target of derogatory assessment and discrimination.

All this may explain some of the troubles we incur as speakers or receivers of generic statements about genders. What is explained, though, it not avoided or overcome. If Leslie is right about the way in which generics function, and therefore about how gender concepts work, what should philosophers do with generalizations about social kinds, gender in particular? Should we avoid using gender generics (for example, assertions the subject of which is expressed by bare plurals like “men” and “women”)? Leslie (2017) recommends the use of adjectives in place of nouns, since the former do not activate essentialist beliefs to the same extent as the latter: but this would hardly work with gender, and might even be a waste of energy, distracting us from tackling other sources of gender discrimination (as Saul 2017 warns). Should we, then, examine and reject the gender generics which fail to meet the conditions for a generic statement to be true (for example, the properties they predicate are not appropriate to generalization – not really “striking”, not really functional to the primary social function of the category)? Or would it be enough to argue that the normative understanding of gender generics is ill-founded? Establishing this last point would cancel part of the harmful implications of the use of gender generics, but would still not be enough to stop them being used to discriminate a whole gender category.

At this point, there is a side issue that needs to be mentioned: I have assumed throughout that philosophy has a professional deontology. But does it actually have a deontology and should it have one? And if so, what kind? Some colleagues might say that philosophers should only seek truth and give reasons for those claims that they make in so doing. In the 20<sup>th</sup> century, numerous thinkers added to this, or else preferred to say that philosophers should refrain

---

<sup>8</sup> About children’s understanding of generics, see Gelman (2010); Gelman, Goetz, Sarnecka and Flukes (2008); Gelman and Raman (2003); Hollander, Gelman and Star (2002); Leslie (2012). About recalling quantified statements as generics, see Gelman, Sánchez Tapia and Leslie (2016); Leslie and Gelman (2012).

<sup>9</sup> Gelman, Ware and Kleinberg (2010); Rhodes, Leslie and Tworek (2012).

from making claims that are not suitable for being assessed as true or false, and should help non-philosophers distinguish claims that are so suitable from claims that are not. But apart from this direct or indirect relationship to truth, and the appeal to reasons, do philosophers have any further responsibilities as regards the society to which they belong? My reply to this question is simple: everybody is responsible for what her speech brings about (illocutions) and to some extent also for their consequences on psychological attitudes and behavior (perlocutions); but even more responsible is whoever knows more or sees farther than others, as should be true of philosophers, or is in a position that can *de facto* have influence on others, as sometimes happens to them. I therefore think that philosophers facing choices about how to deal with matters that have moral or political implications should care as much about what is right as about what is true. Incidentally, these aims may well not be in competition. On the negative side, a philosopher should avoid licensing implications that may be expected to be harmful. On the positive side, a number of morally or politically relevant goals, such as freedom of individual realization, self-determination (within reasonable limits), smoothness in interpersonal relations, and transparency and reversibility of relational asymmetries, are such that philosophy can and should help in the pursuit of. Returning in the light of these reflections to Leslie's worries and recommendations about the harmful assumptions associated with social kind and gender generics, I think that they are basically legitimate and worth consideration. Obviously, changing people's linguistic habits by decree is obviously not a viable option, and after all, it is doubtful whether there is any simple way in which it would be worth changing them: but research on how certain harmful assumptions about genders are communicated and how their communication could be avoided or their content questioned can certainly be of some help on our way towards dealing with gender concepts in correct and appropriate ways.

**4. Gender issues and intersubjective recognition**

It seems that we have come to an impasse. We can make socially shared assumptions about gender explicit and criticize these assumptions. We can criticize not just their content, but also their essentialist implications. We are aware that certain quite common ways of speaking (the use of gender generics), along with their important cognitive function, may also carry harmful presuppositions about gender categories, play a normative function, and pave the way to discrimination. But all this does not seem to give firm grounds to the rejection of gender discrimination, the refutation of essentialist beliefs about the genders, and the neutralization of the normative import of gender concepts. To these aims, we might profit from another (albeit converging) line of reasoning: I would suggest that philosophy tackled gender issues by considering them as a test of people's capability and willingness to recognize others as subjects or, if you prefer, as human persons.<sup>10</sup>

By "intersubjective recognition", I am referring to the process by which two human individuals recognize each other as subjects. It is a basic process in human life, integral both to interpersonal relations and to broader social ones.<sup>11</sup> It is not a unilateral process. An attitude attributing subjectivity to an individual is not a private fact about the recognizer but, through

---

<sup>10</sup> In what follows I will speak of "subjects", in order to leave issues about the definition of "person" aside.

<sup>11</sup> Many hints at this process and its indispensable role can be found in analytic philosophy. Recognition of mental functions, capacities and dispositions is taken into consideration and even given a central place in accounts provided by various philosophers of the process of understanding another subject's language and meaning. At the very least, I would like to mention Davidson's Charity Principle (1984), Grice's Cooperative Principle (1989: 22-40), and Dennett's "intentional stance" (1987). I have touched upon this theme on various occasions, most recently in my (2013). The idea of intersubjective recognition would also profit from a comparison with the theme of intersubjectivity in phenomenology, but the task is so huge that I am in no position to even begin hinting at it.

her behavior and speech, contributes to establishing the intersubjectively shared space of the encounter and to characterizing its participants as subjects endowed with certain attitudes, dispositions, and other action-related properties. It cannot in general be entertained for long if there is no response or feedback on the part of the recognition's target. The recognition of the other's subjectivity, therefore, necessarily involves a framework of intersubjectivity.

Being incapable of such recognition (which involves, for example, not realising that beliefs, intentions, desires and other attitudes vary across individuals and thus being unable to correctly attribute attitudes to others) is not merely being insensitive, but is actually a symptom of certain kinds of psychiatric disorders (see Baron Cohen 1997, Frith 2003). Local, but systematic, refusal to recognize certain others as subjects, possibly associated with the refusal of granting them human rights – as happens in reduction to slavery, or in genocide – is a unanimously condemned stance. Indeed, it should be pointed out that those who (for some reason of their own) do not want to condemn certain other people for taking such a stance make no attempt to defend them directly, but adopt a negationist strategy instead, attempting to deny that the events manifesting that stance have taken place or that they were an actual manifestation of that stance. There is no doubt that a responsible philosophy, in analysing subject-other relations, should foster practices of intersubjective recognition, rather than hinder them, and the first thing to do is to pay attention to how they work.

Recognizing someone as a subject involves assigning her various kinds of competence: attitudes, dispositions, but also rights, obligations and other deontic properties (amounting to what may be called, adapting a notion from narrative semiotics, the subject's "deontic competence")<sup>12</sup> that sometimes correspond to some kind of specific social status. It is tempting to think of this recognition as an all-in-one attribution, but while that may be so by default, in cases in which there are hurdles or problems either on the part of the recognizing subject or on the part of the subject to be recognized, in order to understand what is going on it is necessary to analyze various levels of attribution that may not all be active at the same time. Recognition of the other as a subject, then, consists at the very least of recognizing her as a center of perspective, while further steps consist of recognizing her as capable of (or actually exercising) agency, as a speaker mastering a language and expressing meaning by it, and finally, as a speaker-agent of illocutionary acts, endowed with the deontic competence necessary for her designed performance, and affecting, by her acts, the deontic competence of addressees or other target participants. Recognition of another human being as a subject is completely achieved when it involves all these levels of subjectivity and moreover, is granted unconditionally. Indeed, it can only be flawed when given under-a-condition: for example, if I take you to be an agent only insofar as you do certain things I want to be done, or to be a speaker insofar as that is indispensable to the exercise of a certain subordinate social role, or to be a center of perspective only when you perform your gender role in a way functional to my needs. When recognition is conditional in these ways, it can be suddenly withdrawn in ways which show that it was never actual and effective. It becomes apparent that I never really expected you to have your own perspective on your relationship to me, never really thought you were capable of autonomous decision-making, and never would grant that you be in a position to perform authority-presupposing illocutionary acts.

---

<sup>12</sup> In narrative semiotics (Greimas 1970, 1983; Greimas and Courtés 1979), the dispositions and the deontic properties that prompt the subject's action or enable her to act constitute her "modal competence" for acting. Focusing upon an agent's deontic properties, which in my perspective differ from attitudes or dispositions because they are granted and withdrawn on the basis of social or interpersonal agreement, I use the label "deontic competence". In my work on illocutionary acts since the 1980s, I have connected changes in the deontic competence of participants to illocutionary acts as their conventional effects (1984, 1989).

When the other is recognized as a subject because he or she is similar to us, we are faced with an ambiguous kind of recognition. A recognition based upon the assumption that the subject recognized is similar to the recognizer would indeed be a conditional one and would therefore be flawed. But what if the cognitive bases of recognition, consisting in an ability to attribute attitudes and agency to other people, should involve some kind of simulation of the others' stance and perspective (as in simulation theory, see Goldman 2006), on the basis of our own? Would this amount to making every recognition dependent on similarities between the subject to be recognized and the recognizing one? In many cases, perhaps, it will be enough if the interlocutors' capacities to entertain attitudes and to act are considered similar, or even just parallel. Moreover, if it is true, as research on the role of mirror neurons in social cognition suggests (see Rizzolatti and Craighero 2004, Rizzolatti and Sinigaglia 2007), that what we do is simply use the same action schemes both to project our own action and to understand what another agent is doing, this clearly does not require any previous assessment of similarities between recognizer and recognized. Rather, the feeling of a common belonging to humankind (or to the animal kingdom, or to life in the universe) will come later. But if there is a category of which both the recognizing and the recognized subject are held to be members, and this membership (involving participation in the category's "essence") works as a condition for the recognition, that recognition is flawed.

I would say that women have long suffered from being recognized as subjects both conditionally and partially (namely, not at all levels). This may happen to men too, in certain circumstances and for specific motivations; with racism, by the way, it happens to both genders at once; but for women, it is a chronic condition. The "silencing" of women, so much discussed in the philosophy of language and in political philosophy (Langton 1993 and Hornsby and Langton 1998 are key contributions to this debate),<sup>13</sup> can be seen as a manifestation of the failure to recognize the other's subjectivity and deontic competence unconditionally: a typical and often dramatic case of "silencing" occurs when a woman refuses a man's sexual advances, but the force of her utterance and the attitudes and the deontic competence it involves, both as regards the preconditions and the result of her speech act, are not recognized. It is no coincidence that the women's movement of the 1970s based feminine autonomy upon unconditional intersubjective recognition occurring among women: the words used for describing this process were less abstract and more emotional, but insofar as it worked, that was what was going on.

The issue of intersubjective recognition could be further investigated by considering it in the narrative dimension. I am not referring here to narrative theories of the subject in the philosophy of mind, but to the analytical techniques and terminology of narrative semiotics (Greimas 1970, 1983; Greimas and Courtés 1979). There, the Subject is first of all a narrative "actant", that is, a narrative element defined for its functions in the narrative syntax: receiving a task, carrying it out, and being rewarded. The Subject is then enriched by the attribution of dispositions and deontic properties, to be turned into a more specific "actantial role", and is manifested in discourse by an agent (an "actor") endowed with descriptive characteristics. Whatever the details of the analytic terminology employed, in narrative semiotics the role of Subject is not one that can be played in isolation: another function in the narrative syntax is required to be at work in order for the Subject to receive from it its task, as well as its reward. That function, called the "Destinant", is therefore indispensable to there being a Subject at all. It is the Destinant who assigns the Subject his or her task and

---

<sup>13</sup> Contributions to the debate include: Bird (2002); Caponetto (2017); Hornsby (2011); Jacobson (1995); Maitra (2009); McGowan (2009), (2017); McGowan et al. (2011); Mikkola (2011); Sbisà (2009).

after the performance, recognizes the Subject as the performer. In the terms used here, the Subject of a narrative exists as such insofar as its Destinant recognizes it. This confirms that our understanding of what it is to be a Subject (insofar as the structure of narratives reflects it) involves recognition of subjectivity as an indispensable part of the picture. A problem is raised, though, by the fact that the standard Destinant-Subject relationship appears to be mono-directional and hierarchical and therefore not suited to account for the intersubjective character of recognition of subjectivity. Not only are the functions of Destinant and Subject such that it is the Subject that depends on the Destinant, not the converse, but also, this functional asymmetry is typically reflected in the descriptive characteristics that Destinant and Subject are endowed with: the Destinant, which is also the source of the values of the narrative, does not belong to the same environment as the Subject and is therefore presented as “transcendent” with respect to the Subject and its world. But intersubjective recognition, as I have presented it here, is realized completely only if it is bi-directional. The kind of narrative underlying actual intersubjective recognition should therefore envisage a Destinant whose manipulation and assessment activities are carried out in the same environment as the Subject’s performance, so that in principle agents expressing Destinant and Subject can switch their roles. Not by chance, it is especially in narratives in which the role of Subject is played by a female agent that this non-transcendence of the Destinant manifests itself: while on the one hand, this confirms that the feminine gender has no access to transcendent sources of values, on the other, it opens up the relationship of Subject and Destinant to reversibility (on this, see Sbisà 2017, 41-42) and to the bi-directionality of intersubjective recognition.

To sum up, my suggestion is that further reflection on the ways in which intersubjective recognition is achieved and made manifest (or on the ways in which it fails and is revealed as flawed) can throw light on the dynamics of human relations, both positive and negative, and among these, on gender relations.<sup>14</sup> This reflection could also elicit an increased awareness of why gender discrimination is ill-founded and should be stopped, why essentialist beliefs about the genders should be abandoned, and why gender concepts should not be granted normativity. In fact, discrimination is a process by which a subject is denied full recognition because of his or her belonging, or not belonging, to a certain group. Essentialist beliefs pave the way to discrimination or at any rate, make recognition conditional on the subject having a certain essence. Lastly, the normative function of gender concepts might turn out to arise from the pressure put on the subject seeking recognition by the conditions on which the alleged recognition would be granted.

#### REFERENCES

- Antony, L. (2012). Different Voices or Perfect Storm: Why Are There So Few Women in Philosophy?. *Journal of Social Philosophy* 43 (3), 227–255;
- Baron, B. & Kotthoff, H. (2001) (eds), *Gender in interaction: perspectives on femininity and masculinity in ethnography and discourse*. Amsterdam: John Benjamins;
- Baron-Cohen, S. (1997). *Mindblindness. An essay on autism and theory of mind*. Cambridge, Mass: the MIT Press;
- Bird, A. (2002). Illocutionary silencing. *Pacific Philosophical Quarterly* 83, 1-15;
- Braidotti, R. (1991). *Patterns of Dissonance*, Cambridge: Polity Press;

---

<sup>14</sup> Notwithstanding the great differences in the terms in which the two proposals are expressed, there may be similarities or even some complementarity between tackling gender issues in terms of recognition of subjectivity or failure thereof, as I have been suggesting to do in this section, and Mikkola’s humanist feminism as outlined in her work on dehumanization (Mikkola 2016).



- Butler, J. (1999). *Gender Trouble. Feminism and the subversion of identity*. New York: Routledge. 1<sup>st</sup> ed. 1990;
- Caponetto, L. (2017). On Silencing, Authority, and the Act of Refusal. *Rivista di estetica* 64, 35-52;
- Davidson, D. (1984). Radical interpretation. In D. Davidson (1984), *Inquiries into Truth and Interpretation*, Oxford: Oxford University Press, 125-139;
- Dennett, D. C. (1987). *The intentional stance*. Cambridge, Mass: the MIT Press;
- Fairclough, N. (1995). *Critical discourse analysis. The critical study of language*. London: Longman;
- Fricke, M. & Hornsby, J. (2000) (eds), *The Cambridge companion to feminism in philosophy*. Cambridge: Cambridge University Press;
- Frith, U. (2003). *Autism: explaining the enigma*. Oxford: Blackwell. 1<sup>st</sup> ed. 1991;
- Gelman, S. A. (2010). Generics as a Window onto Young Children's Concepts. In F.J. Pelletier (ed.) (2010), *Kinds, Things, and Stuff: The Cognitive Side of Generics and Mass Terms* (New Directions in Cognitive Science v. 12), New York: Oxford University Press, 100-122;
- Gelman, S. A., Goetz, P.J., Sarnecka, B.S. & Flukes, J. (2008). Generic Language in Parent-Child Conversations. *Language Learning and Development*, 4(1), 1-31;
- Gelman, S. A. & Raman, L. (2003). Preschool Children Use Linguistic Form Class and Pragmatic Cues to Interpret Generics. *Child Development* 74(1), 308-325;
- Gelman, S. A., Sánchez Tapia, I. & Leslie, S. J. (2016). Memory for Generic and Quantified Sentences in Spanish-speaking Children and Adults. *Journal of Child Language* 43(6), 1231-44;
- Gelman, S. A., Ware, E. A. & Kleinberg, F. (2010). Effects of Generic Language on Category Content and Structure. *Cognitive Psychology*, 61 (3), 273-301;
- Gilligan, C. (1982). *In a different voice*. Cambridge, Mass: Harvard University Press;
- Goldman, A. (2006). *Simulating Minds: The Philosophy, Psychology, and Neuroscience of Mindreading*. Oxford: Oxford University Press;
- Greimas, A. J. (1970). *Du Sens*. Paris, Seuil;
- Greimas, A. J. (1983). *Du Sens II*. Paris, Seuil;
- Greimas, A. J. & Courtés, J. (1979). *Sémiotique. Dictionnaire raisonné de la théorie du langage*, Paris: Seuil. Engl. transl. by L. Crist et al., *Semiotics and language: an analytical dictionary*, 1982, Bloomington: Indiana University Press;
- Grice, P. (1989). *Studies in the Way of Words*. Cambridge, Mass: Harvard University Press;
- Heinämaa, S. (2014). 'An equivocal couple overwhelmed with life': A phenomenological analysis of pregnancy, *philoSOPHIA* 4 (1), 12-49;
- Hollander, M. A., Gelman, S.A. & Star, J. (2002). Children's Interpretation of Generic Noun Phrases. *Developmental Psychology*, 38(6), 883-894;
- Hornsby, J. (1994). Illocution and its significance. In S. L. Tsohatzidis (1994) (ed.), *Foundations of Speech Act Theory*, London: Routledge, 187-207;
- Hornsby, J. (2011). Subordination, silencing, and two ideas of illocution. *Jurispudence* 2, 379-385;
- Hornsby, J. & Langton, R. (1998). Free Speech and Illocution. *Legal Theory* 4, 21-37;
- Irigaray, L. (1974). *Speculum. De l'autre femme*. Paris: Minuit;
- Irigaray, L. (1977). *Ce sexe qui n'en est pas un*. Paris: Minuit;
- Jacobson, D. (1995). Freedom of speech acts? A response to Langton. *Philosophy & Public Affairs*, 24, 64-79;
- Kotthoff, H. & Wodak, R. (1997) (eds), *Communicating Gender in Context*. Amsterdam: John Benjamins;
- Lakoff, R. T. (1975). *Language and Woman's Place*. New York: Harper and Row;
- Langton, R. (1993). Speech Acts and Unspeakable Acts. *Philosophy & Public Affairs* 22(4), 292-330;
- Langton, R. & West, C. (1999). Scorekeeping in a Pornographic Language Game. *Australasian Journal of Philosophy* 77(3), 303-319;

- Leslie, S. J. (2007). Generics and the Structure of the Mind. *Philosophical Perspectives* 21(1), 375–403;
- Leslie, S. J. (2008). Generics: Cognition and Acquisition. *Philosophical Review*, 117(1), 1–47;
- Leslie, S. J. (2012). Generics articulate default generalizations. *Revue Linguistique de Vincennes* 41, 25–44;
- Leslie, S. J. (2017). The original sin of cognition. *Journal of Philosophy* 114 (8), 393–421;
- Leslie, S. J. & Gelman, S. A. (2012). Quantified Statements Are Recalled as Generics: Evidence from Preschool Children and Adults. *Cognitive Psychology*, 64 (3), 186–214;
- Leslie, S. J. & Lerner, A. (2016). Generic Generalizations. In *The Stanford Encyclopedia of Philosophy* (Winter 2016 Edition), E. N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/win2016/entries/generics/>;
- Lorber, J. & Farrell, S. A. (1991) (eds). *The social construction of gender*. Newbury: Sage;
- Maitra, I. (2009). Silencing Speech. *Canadian Journal of Philosophy*, 39 (2), 309–338;
- McGowan, M. K. (2009). On silencing and sexual refusal. *The Journal of Political Philosophy*, 17 (4), 487–494;
- McGowan, M. K. (2017). On multiple types of silencing. In M. Mikkola, *Beyond Speech. Pornography and Analytic Feminist Philosophy*, Oxford: Oxford University Press, 39–57;
- McGowan, M. K., Helmers, S., Stolzenberg, J. & Adelman, A. (2011). A partial defense of illocutionary silencing. *Hypathia* 26 (1), 132–149;
- Mikkola, M. (2011). Illocution, silencing and the act of refusal. *Pacific Philosophical Quarterly* 92, 415–437;
- Mikkola, M. (2016). *The Wrong of Injustice: Dehumanization and its Role in Feminist Philosophy*. Oxford: Oxford University Press;
- Rizzolatti, G. & L. Craighero, L. (2004). The Mirror-Neuron System. *Annual Review of Neuroscience*, 27, 169–92;
- Rizzolatti, G. & Sinigaglia, C. (2007). Mirror neurons and motor intentionality. *Functional Neurology*, 22(4), 205–210;
- Rhodes, M., Leslie, S. J. & Tworek, C. (2012). Cultural Transmission of Social Essentialism. *Proceedings of the National Academy of Sciences*, 59 (34), 13526–31;
- Saul, J. M. (2003). *Feminism. Issues and Arguments*. Oxford: Oxford University Press;
- Saul, J. M. (2017). Are generics especially pernicious?. *Inquiry*, DOI: 10.1080/0020174X.2017.1285995;
- Sbisà, M. (1976). Speech acts e femminilità. *Problemi*, 47, 260–83;
- Sbisà, M. (1984a). On illocutionary types. *Journal of Pragmatics*, 8, 93–112;
- Sbisà, M. (1984b). *La mamma di carta. Per una critica dello stereotipo materno*. Milano: Emme;
- Sbisà, M. (1989). *Linguaggio, ragione, interazione*. Bologna: Il Mulino. 2nd edition (as an E-book), 2009, Trieste: EUT;
- Sbisà, M. (1996). Feminine subject and female body in discourse about childbirth. *European Journal of Women's Studies* 3, 1996, 363–76;
- Sbisà, M. (2009). Illocution and silencing. In B. Fraser & K. Turner (eds), *Language in Life, and a Life in Language: Jacob Mey - A Festschrift*, Bradford: Emerald, 351–357;
- Sbisà, M. (2013). Soggetto e riconoscimento. In M. Leone & I. Pezzini (eds), *Semiotica delle soggettività. Per Omar*, Roma: Aracne, 169–192;
- Sbisà, M. (2017). Il soggetto femminile: dimensioni di analisi. In L. Borghi & R. Svandrlik (eds) *S/Oggetti Immaginari. Letterature comparate al femminile*, Ebook@Women, 36–42. 1st ed. 1996;
- Williamson, T. (2006). Past the Linguistic Turn?. In B. Leiter (ed.). *The future for philosophy*. Oxford: Oxford University Press, 106–128;
- Williamson, T. (2007). *The Philosophy of Philosophy*. Oxford: Blackwell;
- Wittgenstein, L. (1922). *Tractatus Logico-Philosophicus* (with English translation). London:

Routledge and Kegan Paul;

Wittgenstein, L. (1953). *Philosophische Untersuchungen/Philosophical Investigations* (with English translation). Oxford: Blackwell;

Wodak, R. & Chilton, P. (2005). *A New Agenda in (Critical) Discourse Analysis. Theory, methodology and interdisciplinarity*. Amsterdam: John Benjamins.

---

LAURA CAPONETTO

Università Vita-Salute San Raffaele  
lauracaponetto@gmail.com

---

# FILOSOFIA DEL LINGUAGGIO FEMMINISTA, ATTI LINGUISTICI E RIDUZIONE AL SILENZIO

---

## abstract

*In anni recenti, la teoria degli atti linguistici di Austin è stata impiegata in ambito di filosofia del linguaggio femminista per dar corpo alla nozione di riduzione al silenzio, messa originariamente in campo da MacKinnon nel quadro del dibattito su pornografia e censura. L'idea, sostenuta da più voci, è che certi materiali pornografici contribuiscano a creare un clima comunicativo ostile alle donne, che impedisce loro di compiere certi atti illocutori – primo fra tutti, l'atto del rifiuto di avances sessuali. Nel presente lavoro, metto a confronto due declinazioni della nozione di riduzione al silenzio, offerte rispettivamente da Hornsby e Langton e da McGowan. Offro, inoltre, un'analisi inedita del rifiuto che inficia parzialmente la proposta di McGowan. L'obiettivo è quello di fornire uno spaccato di come la teoria austiniana possa essere adoperata, a mo' di "cassetta degli attrezzi", per portare alla luce forme di ingiustizia discorsiva passate largamente inosservate.*

---

## keywords

*filosofia del linguaggio femminista, atti linguistici, riduzione al silenzio, rifiuto, autorità*

**1. Filosofia del linguaggio femminista: temi e strumenti**

Nell'arco degli ultimi trent'anni, la filosofia del linguaggio femminista ha mosso i primi passi, è cresciuta e si è consolidata in quello che oggi è un settore disciplinare dotato di una fisionomia specifica, al crocevia tra filosofia del linguaggio e studi di genere.

Nelle sue prime fasi, la disciplina (che, strettamente parlando, non poteva ancora essere definita tale, mancando di uno statuto autonomo) assume carattere prevalentemente negativo. Molti lavori iniziali prendono le mosse dalla critica alla presunta neutralità del linguaggio, reinterpretato come deposito di disuguaglianze sociali e asimmetrie di genere. Obiettivo comune di autrici come Deborah Cameron, Janice Moulton e, più recentemente, Penelope Eckert e Sally McConnell-Ginet (per citarne solo alcune) è stato quello di svelare il sessismo inscrito nel linguaggio e quindi reclamarne azioni di riforma. Si pensi alle proposte, avanzate a partire dai primi anni Ottanta, di riforma degli usi solo apparentemente neutrali del pronome maschile ('lui', 'egli') per riferirsi a soggetti di cui non si conosce l'identità di genere o di sostantivi come 'uomo' o 'uomini' per riferirsi agli esseri umani in generale (donne incluse). Si tratta di proposte oggi accolte da buona parte della ricerca scientifica in lingua inglese: l'APA (*American Psychological Association*), per fare un esempio autorevole, ha introdotto delle linee guida per un linguaggio senza pregiudizi ("bias-free"). Tra le "buone pratiche", si raccomanda di evitare l'uso generico dei pronomi di genere 'he' e 'she'; in alternativa, si consiglia di riformulare gli enunciati in oggetto al plurale o sostituire il pronome con un sostantivo come 'person' o 'individual' – laddove ciò non fosse possibile, si ammette l'uso della formula mista 'he or she'. Si suggerisce inoltre l'uso del pronome plurale 'they' per riferirsi a soggetti singolari che preferiscano non essere associati al genere maschile né al genere femminile<sup>1</sup>.

In anni recenti, agli obiettivi critici della prima filosofia del linguaggio femminista si sono affiancati programmi di ricerca positivi. Nella filosofia analitica del linguaggio, in prima battuta criticata poiché ritenuta espressione di una prospettiva teorica maschile e individualista, molte autrici hanno rintracciato risorse utili al progetto femminista e ai suoi scopi di emancipazione. Tra queste, un ruolo chiave è tuttora svolto dalla teoria degli atti linguistici di John Austin (1962), che è stata sfruttata in maniera inedita per gettar luce sulla

---

<sup>1</sup> "Respect people's preferences; call people what they prefer to be called [...]; you may need to ask your participants which designations they prefer, particularly when preferred designations are being debated within groups" (*APA Publication Manual*, 2010, p. 72).

dimensione politica del linguaggio – e su come il linguaggio costituisca uno dei luoghi cardine dell’oppressione. Non solo deposito dunque, ma origine stessa di sessismo e discriminazione<sup>2</sup>. Tra le applicazioni più efficaci (e sorprendenti) della teoria austiniana spicca la difesa della controversa tesi anti-pornografia di Catharine MacKinnon a opera di Jennifer Hornsby e Rae Langton<sup>3</sup>. La pornografia, a detta di MacKinnon (1987, 1993), *riduce al silenzio* (“silences”) le donne e, in tal senso, rappresenta una grave violazione del loro diritto alla libertà di espressione. L’argomento, rapidamente dismissed da alcuni come confuso, indifendibile, se non palesemente falso<sup>4</sup>, è stato riletto da Hornsby e Langton (Hornsby 1993, 2014; Langton 1993, 2009; Hornsby & Langton 1998) alla luce della teoria di Austin – e reso filosoficamente plausibile. Dire che la pornografia riduce al silenzio le donne significa dire, in quest’ottica, che la pornografia svuota certi enunciati proferiti dalle donne del loro potenziale illocutorio. Pur consentendo loro di compiere atti locutori, la pornografia impedirebbe alle donne, in certe circostanze, di “fare cose con le parole”. In primo luogo, come vedremo a breve, rifiutare avances sessuali maschili.

La nozione di *riduzione al silenzio* (“illocutionary silencing”) messa in campo da Hornsby e Langton è stata ripresa ed elaborata da diverse autrici (si vedano, a titolo esemplificativo, Wieland 2007; Maitra 2009; Sbisà 2009; McGowan 2009, 2014) e assume oggi la veste di una nozione-cappello volta a raggruppare diverse tipologie di fallimento illocutorio<sup>5</sup>. Benché l’atto del rifiuto sia da tutte le autrici menzionate individuato come esempio paradigmatico di riduzione al silenzio, ben poco spazio è stato dedicato in letteratura a cosa significhi rifiutare. Che *tipo* di atto è il rifiuto? Quali sono le sue condizioni di successo? Quando dire “no” conta come un rifiuto? Va da sé che la risposta a questi e altri interrogativi simili sia fondamentale al fine di chiarire i meccanismi mediante i quali il rifiuto è o può essere ridotto al silenzio.

Il presente lavoro si ripropone, da un lato, di offrire un’analisi dell’atto del rifiuto e, dall’altro, di valutare una delle più recenti e interessanti riformulazioni della nozione di riduzione al silenzio: quella proposta da Mary Kate McGowan (2009). Così facendo, intendo fornire uno spaccato di come sia possibile adoperare la teoria degli atti linguistici come “cassetta degli attrezzi”, come insieme di strumenti per render conto di certe forme di ingiustizia passate largamente inosservate. In particolare, per render conto di come le parole di alcuni siano ingiustamente e sistematicamente private di forza in virtù della sola identità sociale di chi le proferisce.

L’articolo è strutturato come segue. Dopo aver introdotto la tesi anti-pornografia di MacKinnon (§ 2) e la lettura in chiave austiniana fornita da Hornsby e Langton (§ 2.1), presento la nozione di riduzione al silenzio per com’è stata caratterizzata da McGowan. Nella prospettiva di McGowan, la riduzione al silenzio si lega a doppio filo al mancato riconoscimento dell’autorità del parlante: le donne non riuscirebbero a rifiutare le avances sessuali maschili poiché sistematicamente ritenute prive di autorità (persino sul proprio corpo) (§ 2.2). Passo poi ad analizzare l’atto del rifiuto (§ 3). Come risulterà evidente, il rifiuto è un atto di “secondo turno”, un atto felice solo se eseguito in risposta a un altro

---

2 Per un’introduzione alla filosofia del linguaggio femminista, si veda Saul & Diaz-Leon (2017). Un’accurata indagine sui legami tra linguaggio e genere è invece contenuta in Bianchi (2009).

3 La nozione austiniana di *performatività*, insieme con la distinzione tra illocuzione e perlocuzione, è stata ripresa anche da Judith Butler (1990, 1997) nella sua riflessione sull’identità di genere come costruzione sociale e culturale. Su tale riflessione, che va ben al di là della filosofia del linguaggio, investendo gli ambiti della metafisica dei generi e della filosofia pratica, non ci soffermeremo. Per una discussione critica del pensiero di Butler, si veda Salih (2002).

4 Si vedano, per esempio, le reazioni all’argomento di MacKinnon contenute in Parent (1990) e Dworkin (1991).

5 Per un’elaborazione di quest’idea, si vedano Caponetto (2016) e McGowan (2017).

(una richiesta, un'offerta, un invito) precedentemente compiuto dall'interlocutore (§ 3.1). Il rifiuto si configura come un "atto autoritativo" – ovvero un atto che necessita dell'autorità del parlante sul dominio rilevante – *solo se* eseguito in risposta a una richiesta di permesso. In caso contrario, l'atto può essere felicemente compiuto pur in assenza dell'autorità del parlante (§ 3.2). Sulla scorta dell'analisi condotta, nell'ultima sezione metto in luce due importanti debolezze della proposta di McGowan. La prima è di natura etica: interpretare, come fa McGowan, le avances sessuali maschili come richieste di permesso e i "no" delle donne come atti autoritativi significa presupporre un modello di sessualità "a senso unico", che vede l'uomo come soggetto desiderante e agente e priva ingiustamente le donne di *sexual agency*. La seconda ha invece a che fare con la coerenza interna della proposta: se nel fare avances sessuali gli uomini compiono tipicamente richieste di permesso, allora – avrò modo di spiegare perché – non si dà il caso che questi non riconoscano l'autorità delle donne sul proprio corpo, in contraddizione con quanto sostenuto da McGowan (§ 4).

## 2. Pornografia e atti linguistici

La pornografia<sup>6</sup> non è oscena o sconveniente ai fini del pubblico decoro. Non va osteggiata perché immorale, come si sente spesso reclamare in ambienti conservatori. Piuttosto, va contrastata perché *riduce al silenzio le donne*, ledendo il loro diritto alla libertà di espressione. È questo il noto argomento anti-pornografia elaborato da MacKinnon a metà degli anni Ottanta. L'argomento, lo dicevamo, ha sin da subito attratto numerose critiche, ma anche dato il via a una feconda linea interpretativa inaugurata da Hornsby e Langton con l'intento di affinare la tesi e renderla immune alle critiche di implausibilità. L'interpretazione, di ispirazione austiniiana, si impernia sull'idea che la pornografia riduca (o possa ridurre<sup>7</sup>) al silenzio le donne in un senso ben preciso: ostacolando nel compimento di certi atti illocutori – per esempio, l'atto del rifiuto di avances sessuali<sup>8</sup>. La nozione filosofica chiave è qui quella di illocuzione, che Austin contrappone a locuzione e perlocuzione. Vediamo brevemente cosa si intende per atto illocutorio e quali sono le condizioni da soddisfare perché un atto entri in vigore.

### 2.1. Riduzione al silenzio e uptake

Il nucleo della teoria degli atti linguistici di Austin può essere riassunto nello slogan *proferire enunciati equivale a compiere atti*. L'atto linguistico è per Austin unitario, ma, a seconda dell'aspetto considerato, può essere descritto come atto *locutorio* o atto *di* dire qualcosa (proferimento di un enunciato sintatticamente ben formato e dotato di significato), atto *illocutorio* o atto che si compie *nel* dire qualcosa (azione compiuta attraverso il proferimento dell'enunciato), atto *perlocutorio* o atto che si compie *col* dire qualcosa (conseguenze non convenzionali dell'enunciato proferito). Ogni aspetto dell'atto è soggetto a determinati difetti, che possono causarne il fallimento. Concentriamoci sull'aspetto illocutorio, che è quello più rilevante ai fini della delucidazione della proposta di Hornsby e Langton. Secondo Austin, l'atto illocutorio fallisce (è "nullo") se non sono soddisfatte certe condizioni di

6 Il termine 'pornografia', nell'uso di MacKinnon (1987, p. 176), non si riferisce all'intero insieme dei materiali comunemente etichettati come pornografici, ma a un suo sottoinsieme proprio, formato da tutti e soli quei materiali (tipicamente rivolti a un pubblico maschile eterosessuale) che subordinano le donne, raffigurandole alla stregua di oggetti sessuali e in scenari particolarmente degradanti. Nel corso del presente lavoro, il sostantivo 'pornografia' e i suoi derivati aggettivali vengono sempre usati in questo senso ristretto.

7 L'obiettivo di Hornsby e Langton è quello di mostrare la plausibilità filosofica della tesi di MacKinnon, non la verità della stessa – compito, quest'ultimo, di natura empirica. (Per alcuni risultati sperimentali a parziale supporto della tesi di riduzione al silenzio si vedano Milburn *et al.* (2000) e Malamuth *et al.* (2011).)

8 Che ridurre qualcuno al silenzio illocutorio equivalga a violarne la libertà di espressione è materia di disputa. Per sostenere l'equivalenza occorre infatti impegnarsi a una visione della libertà di espressione come libertà di illocuzione oltretutto di locuzione, che non tutti sono disposti a sottoscrivere. Per una critica di questa concezione "ampia" della libertà di espressione si veda Dworkin (1991); una convincente risposta è contenuta in West (2003).

felicità. In particolare, si richiedono l'esistenza della procedura invocata e l'appropriatezza di persone e circostanze (condizioni che Austin classifica come "tipo A"), oltretutto la correttezza e la completezza dell'esecuzione della procedura ("tipo B"). A queste bisogna aggiungere la ricezione (*uptake*), introdotta come effetto comune a ogni atto illocutorio, ma trattata alla stregua di una condizione di felicità: Austin (1962, pp. 115-116) precisa infatti che un atto illocutorio può dirsi compiuto solo se l'ascoltatore ha effettivamente compreso il significato e la forza della locuzione proferita dal parlante.

Nel quadro delineato da Hornsby e Langton, la pornografia mette sistematicamente a repentaglio l'*uptake* di alcuni atti compiuti dalle donne, condannandoli al fallimento. Nella fattispecie, ostacola la corretta ricezione dei "no" che le donne proferiscono in risposta ad avances sessuali. La tesi si fonda sulla constatazione che spesso, in film e materiali pornografici di vario genere, quelli che cominciano come stupri sfociano poi in atti sessuali dai quali anche la donna trae godimento. Il suo iniziale rifiuto è presentato come una mera messa in scena allestita allo scopo di fingere pudore e, al contempo, accrescere l'eccitazione dell'uomo. Alcuni fruitori di pornografia potrebbero per questo non riconoscere il "no" delle loro potenziali partner come un rifiuto, ma interpretarlo come parte del gioco, come una strategia messa in atto per accrescere il desiderio maschile<sup>9</sup>. Ecco in che senso la donna è ridotta al silenzio illocutorio: può fisicamente dire "no" ma, in dati contesti, dicendo *no* non riuscirà a fare ciò che intende fare (ovvero rifiutare), poiché il suo atto verrà sistematicamente non recepito per com'è inteso<sup>10</sup>.

Prima di passare oltre, una breve sosta sulla condizione di sistematicità. Hornsby e Langton sostengono che certi enunciati proferiti dalle donne siano *sistematicamente* votati al fallimento. La condizione di sistematicità è costitutiva del fenomeno della riduzione al silenzio: si dà riduzione al silenzio solo se le parole di certi individui vengono svuotate di potenziale illocutorio *in maniera non idiosincratca*. Se un parlante proferisce un certo enunciato e il destinatario non ne coglie la forza perché, per esempio, ha le orecchie tappate, è ubriaco, distratto o incompetente nella lingua parlata dall'interlocutore, il fallimento illocutorio che ne consegue non costituirà un'istanza di silencing. Al contrario, se un parlante proferisce un certo enunciato e il destinatario non ne coglie la forza per via di credenze socialmente condivise che ostacolano o impediscono la ricezione, il fallimento illocutorio che ne consegue sarà un'istanza di silencing. Affinché il mancato *uptake* del rifiuto di una donna sia un esempio di riduzione al silenzio non si richiede, dunque, che *tutti* i "no" proferiti dalle donne in risposta ad avances sessuali vadano a vuoto. Piuttosto, si richiede che, quando ciò accade, non sia per ragioni meramente idiosincratciche, ma per via di false credenze e pregiudizi di genere che intralciano la comunicazione, rendendo per le donne particolarmente difficile compiere atti che avrebbero altrimenti compiuto con facilità.

Hornsby e Langton declinano la nozione di *silencing* nei termini di un sistematico fallimento illocutorio costituito dal mancato riconoscimento, da parte del destinatario, della forza con cui l'enunciato è proferito. A questa tipologia di riduzione al silenzio, McGowan ne affianca un'altra, derivante dal mancato riconoscimento, da parte del destinatario, dell'autorità del parlante. Da un lato, la riduzione al silenzio passa attraverso un fallimento nel processo di

## 2.2. Riduzione al silenzio e autorità dei parlanti

---

<sup>9</sup> La tesi è tanto più plausibile quanto più si abbassa l'età dei fruitori. Si veda in merito il report pubblicato nel 2013 dall'UK Office of the Children's Commissioner, che testimonia come la pornografia costituisca una delle principali fonti di "informazione" sul sesso (e su come ci si comporta nei contesti sessuali) per molti teenagers (Coy *et al.* 2013).

<sup>10</sup> Cfr. Langton, 1993, p. 321: "Sometimes 'no', when spoken by a woman, does not count as the act of refusal. The hearer fails to recognize the utterance as a refusal; uptake is not secured [...]. Since illocutionary force depends, in part, on uptake being secured, the woman fails to refuse".



uptake; dall'altro, attraverso un'indebita negazione di autorità. Nella prospettiva di McGowan, la pornografia riduce al silenzio le donne non solo o non tanto ostacolando la ricezione dei "no" che talvolta proferiscono, ma relegandole a creature-oggetto prive di qualsivoglia autorità, compresa quella sul proprio corpo. Così facendo, la pornografia priverebbe le donne dello status necessario per rifiutare. Ma ricostruiamo l'argomento di McGowan un passo per volta.

Innanzitutto, McGowan classifica il rifiuto come un *atto autoritativo* (*authoritative speech act*): un atto per poter compiere il quale il parlante deve godere di autorità su un certo dominio. Immaginiamo che un soldato semplice, rivolgendosi a uno degli ufficiali del proprio reggimento, dica

(1) Al mio tre, aprire il fuoco,

con l'intenzione di impartire un ordine. Immaginiamo, inoltre, che il destinatario riconosca l'intenzione illocutoria con cui (1) è proferito. In tal caso, benché l'uptake sia assicurato, l'atto "farà cilecca" (per usare il vocabolario austiniano), poiché il soldato non gode dell'autorità necessaria per ordinare all'ufficiale alcunché. Il tentativo del soldato viola una delle condizioni di felicità necessarie individuate da Austin: quella che prescrive l'appropriatezza di chi invoca la procedura in oggetto. Nel contesto immaginato, il soldato non è la persona appropriata per invocare la procedura dell'ordine. Ora, il rifiuto è, a detta di McGowan, simile a un ordine nella misura in cui può essere felicemente eseguito solo da chi abbia autorità sul dominio rilevante. Nel caso del rifiuto di avances sessuali, è necessario che il parlante abbia autorità sul proprio corpo. Chiaramente, ogni essere umano – *qua* persona – detiene un'autorità siffatta; tuttavia, certi uomini sembrano non ascriverne il possesso alle donne (o ad alcune donne). I materiali pornografici, continua McGowan, sono almeno in parte responsabili per questo: raffigurando le donne come meri strumenti per il piacere sessuale maschile, diffondono e alimentano l'odioso pregiudizio secondo cui i corpi delle donne sono a disposizione degli uomini, impedendo così alle prime l'esercizio di certe forme (discorsive e non) di autorità pratica<sup>11</sup>. In ciò che segue, offrirò un'analisi dettagliata dell'atto del rifiuto. Il rifiuto verrà definito come un atto illocutorio di "secondo turno", le cui caratteristiche – ivi inclusa l'autorità del parlante – variano col variare degli atti di "primo turno" cui il rifiuto, di volta in volta, risponde. Per valutare la coerenza filosofica della proposta di McGowan (secondo cui il rifiuto di avances sessuali costituisce un atto autoritativo) bisognerà allora chiarire, in sede preliminare, la natura illocutoria di un'avance sessuale tipo.

**3. Il rifiuto** Il rifiuto non compare tra gli atti illocutori classificati da Austin. Compare però nella tassonomia searliana. Scrivono in merito Searle e Vanderveken (1985, p. 195):

A refusal is the illocutionary denegation of an acceptance [...]. Strictly speaking, one can only accept or refuse a speech act that allows for the option of acceptance or refusal.

Il rifiuto – come l'accettazione, sua controparte illocutoria – non può essere compiuto in isolamento: perché un certo enunciato conti come un rifiuto, deve essere proferito in risposta a un atto (di natura interrogativa) precedentemente compiuto dal destinatario. Il rifiuto

<sup>11</sup> Cfr. McGowan, 2009, p. 492: "A woman says 'No' in response to sexual advances intending to refuse; the man recognizes her intention to refuse, but he falsely believes that she does not have the authority to do so. When this happens, she is silenced in my alternative sense". Una posizione analoga è adottata da Sbisà (2009).

è un'ilocuzione di “secondo turno” o, nel linguaggio dei teorici della conversazione, la seconda parte di una “coppia adiacente” (*adjacency pair*). Per “coppie adiacenti” si intendono quelle sequenze conversazionali costituite da due turni di parola adiacenti ordinati secondo un principio di rilevanza condizionale – ovvero tali che, dato il proferimento di un primo enunciato da parte di A, si renda rilevante o pertinente il proferimento di un secondo enunciato da parte di B (cfr. Levinson 1983, p. 306; Schegloff & Sacks 1973). Esempi di coppie adiacenti sono *domanda-risposta*, *saluto-saluto*, *lamentele-scuse*, e così via. Si noti che il rifiuto è una risposta appropriata<sup>12</sup> a diversi atti di primo turno: *offerte*, *inviti*, *proposte*, *richieste*. Al fine di tratteggiare un quadro completo di cosa significhi rifiutare, diamo una rapida occhiata alle possibili prime parti della coppia ...-rifiuto.

Se vogliamo far fare qualcosa a qualcuno, i mezzi illocutori a nostra disposizione sono molteplici. Esiste infatti un'ampia gamma di atti linguistici (cosiddetti *direttivi*, il cui scopo è quello di orientare la condotta altrui<sup>13</sup>. La scelta del mezzo è funzione del contesto. In certi contesti, possiamo apertamente ordinare a qualcuno di fare qualcosa o esigere che lo faccia; in altri, è invece appropriato proporre al nostro interlocutore di fare ciò che auspichiamo o chiedergli di farlo. *Ordini*, *pretese*, *proposte*, *richieste* sono tutte istanze di ciò che Mark Lance e Rebecca Kukla definiscono “appelli” (“calls”): atti illocutori alla seconda persona singolare o plurale che invitano una specifica risposta da parte del destinatario (o dei destinatari) (Lance & Kukla 2013). Gli appelli possono essere suddivisi in due macro-categorie: *appelli aperti* (o *interrogativi*), come offerte, inviti, proposte, richieste; e *appelli chiusi* (o *imperativi*), come comandi, ordini, pretese, e così via. L'output normativo degli atti afferenti all'una o all'altra categoria è profondamente diverso. Gli appelli chiusi obbligano il destinatario a fare ciò che gli è stato detto di fare; gli appelli aperti, invece, offrono al destinatario una ragione per agire, ma tale ragione non si configura come un obbligo. Ne segue che, a differenza degli appelli chiusi, gli appelli aperti lasciano al ricevente la possibilità di *rifiutare*. Tutti gli appelli invitano risposte comportamentali, ma solo gli appelli aperti invitano una risposta inerentemente illocutoria (di tipo accettazione o rifiuto). Per riagganciarci a quanto detto nel paragrafo precedente, gli appelli aperti costituiscono prime parti di coppie adiacenti, gli appelli chiusi no. Dato un ordine o un comando (felicitamente compiuti), nessuna risposta illocutoria è attesa o pertinente. L'unica risposta adeguata è l'obbedienza<sup>14</sup>.

Tipicamente, quando A proferisce un appello aperto, B non necessita di alcun tipo di autorità per rispondere come meglio crede. Supponiamo che il mio coinquilino mi chieda di lavare la pila di piatti ammonticchiati nel lavello. Per accettare o rifiutare non serve che io occupi alcuna posizione di autorità: posso farlo semplicemente in quanto destinataria dell'appello. Ciò vale anche nei casi in cui la relazione tra i parlanti è asimmetrica. Supponiamo che, nella Virginia di inizio Ottocento, Mr Woodall chieda a uno dei suoi schiavi di svolgere una certa mansione, precisandogli che si tratta di una mera richiesta e che può pertanto dire *no* senza temere ritorsioni. Se il padrone *chiede* (anziché *ordinare* per via indiretta) e rende le sue intenzioni sufficientemente manifeste, allora lo schiavo sarà libero di rifiutare se pur privo di

### 3.1. “Appelli” e atti linguistici di primo turno

### 3.2. Richieste di permesso e altri appelli “aperti”

---

<sup>12</sup> Il rifiuto è una risposta appropriata ma *dispreferita* a diversi atti di primo turno. Per la distinzione tra seconde parti preferite e dispreferite si veda Levinson (1983, p. 307).

<sup>13</sup> La classe dei direttivi è stata introdotta da John Searle e raggruppa buona parte degli esercitivi austiniiani non istituzionali. Cfr. Searle (1975).

<sup>14</sup> Uno può, chiaramente, rifiutarsi di obbedire. Dal punto di vista illocutorio, però, questo non costituisce un rifiuto, ma un annuncio di disobbedienza (un espressivo del tipo ‘Non ho intenzione di fare  $\phi$ ’). Alcune considerazioni in merito si trovano in Searle & Vanderveken (1985, p. 195).

qualsivoglia autorità sul suo interlocutore<sup>15</sup>. Simili considerazioni valgono per *proposte*, *offerte*, *inviti*: un parlante può accettarli o rifiutarli (esibendo, per motivi di cortesia, le ragioni del rifiuto) se è il destinatario dell'appello, ovvero soddisfa una mera condizione di pertinenza.

*Condizione di pertinenza*: un parlante è in diritto di accettare o rifiutare un appello aperto se e solo se ne è il destinatario (o ha ricevuto il beneplacito da parte del destinatario).

Ma questa è solo parte della storia. Ci sono infatti appelli aperti ai quali è possibile rispondere adeguatamente solo se si gode di una qualche forma di autorità. Si considerino gli enunciati che seguono:

- (2) Lavi i piatti, per favore?
- (3) Posso usare il tuo pc?

Nel proferire (2), A sta chiedendo a B di fare  $\varphi$ ; nel proferire (3), A sta chiedendo a B il permesso per fare  $\varphi$ . La prima è una *richiesta semplice*, la seconda è invece una *richiesta di permesso*. Si noti che, nel caso delle richieste semplici,  $\varphi$  non ricade sotto la giurisdizione di B (possiamo ipotizzare che A, in quanto coinquilino di B, abbia uguale accesso alla pila di piatti nel lavello); inoltre, in caso di risposta affermativa, l'azione rappresentata nella proposizione verrà svolta dal destinatario (B). Le richieste semplici sono direttivi searliani in senso proprio: atti con i quali il parlante tenta di far fare qualcosa al destinatario (nella fattispecie, lavare i piatti). Diverso è il caso delle richieste di permesso, che sono tali poiché  $\varphi$  ricade sotto la giurisdizione di B (nel nostro esempio, il computer è di proprietà di B e non di A); inoltre, in caso di risposta affermativa, l'azione rappresentata nella proposizione verrà svolta dal parlante (A). Le richieste di permesso sono direttivi *sui generis*, con cui il parlante tenta di far sì che il destinatario gli accordi un permesso affinché lui stesso possa in seguito fare qualcosa (si veda in merito Cowart 2004). Perché B possa felicemente rispondere "Sì" o "No" a (3), deve sì essere il destinatario dell'appello ma anche l'effettivo proprietario del pc. Supponiamo che A rivolga (3) a B indicando un pc che crede sia di B, ma che in verità è di proprietà di C. In una situazione siffatta, B non sarebbe legittimato a rispondere dal momento che il pc non ricadrebbe sotto la sua giurisdizione. Quando si tratta di rispondere a richieste di permesso, sembra allora che alla condizione di pertinenza sia da affiancare una condizione di autorità.

*Condizione di autorità*: un parlante è in diritto di accettare o rifiutare una richiesta di permesso solo se detiene autorità sul dominio rilevante (o è stato autorizzato a farlo da chi la detiene).

Facciamo il punto. Il compimento di un atto di rifiuto è vincolato al previo proferimento, da parte del destinatario, di un appello "aperto". Se l'appello è una richiesta semplice (o un atto simile, come una proposta, un invito o un'offerta), allora la felicità del rifiuto sarà subordinata al soddisfacimento di una semplice condizione di pertinenza. Se invece l'appello è una richiesta di permesso, la felicità del rifiuto dipenderà dal ricadere o meno dell'attività in

---

<sup>15</sup> Si vedano Kukla (2014, p. 455) e Green (2016) per la discussione di alcuni casi di *amplificazione illocutoria* ("illocutionary amplification") o silencing alla rovescia: casi in cui la posizione sociale dominante occupata dal parlante conferisce alle sue parole una forza maggiore rispetto a quella intesa. Nel nostro esempio, la posizione occupata da Mr Woodall potrebbe far sì che le sue parole vengano sistematicamente recepite come ordini dallo schiavo, anche quando sono intese come richieste e il parlante fa tutto ciò che è in suo potere per rendere manifesta la propria intenzione illocutoria.

oggetto sotto la giurisdizione di chi rifiuta. Le condizioni di felicità del rifiuto variano dunque col variare dell'atto di primo turno cui il rifiuto risponde. In particolare, il rifiuto si configura come un atto autoritativo solo se preceduto da una richiesta di permesso.

Ritorniamo al caso del rifiuto di avances sessuali e alla proposta di McGowan circa la riduzione al silenzio delle donne. Come abbiamo visto nel § 2.2, McGowan classifica il rifiuto come un atto autoritativo e sostiene che ad alcune donne sia negata la possibilità di rifiutare avances sessuali, poiché, a monte, è loro negata autorità sul proprio corpo. A sostegno del ruolo che l'autorità svolge ai fini della felicità del rifiuto (e del rifiuto di avances sessuali in particolare), McGowan presenta il seguente scenario ipotetico. Immaginiamo che Sally dica a Carl che Cindy non è disposta ad avere una relazione sessuale con lui. Anche nel caso in cui Sally convinca Carl a desistere, le sue parole non conteranno come un rifiuto. Diversamente, se è Cindy a comunicare a Carl di non voler fare sesso con lui, le sue parole contano come un chiaro rifiuto. Nella prospettiva di McGowan, quest'asimmetria si spiega alla luce del fatto che il rifiuto è un atto autoritativo – e Cindy possiede sul proprio corpo un'autorità che Sally, com'è ovvio, non ha.

#### 4. Alcune difficoltà

When Cindy says 'No' in response to Carl's sexual advances, Cindy sexually refuses exactly because she thereby denies Carl *permission to proceed*. Having authority over who has sexual access to her body, Cindy is here exercising that authority [...]. Sally cannot refuse on Cindy's behalf exactly because Sally lacks the requisite authority (McGowan 2009, p. 489, corsivo mio).

Da questo passaggio si evince come McGowan interpreti le avances sessuali come richieste di permesso – più precisamente, come richieste di accesso al corpo della persona cui ci si sta rivolgendo. Ciò è in linea con l'analisi condotta nel paragrafo precedente: un rifiuto necessita dell'autorità del parlante solo se preceduto da una richiesta di permesso. Ciononostante, sosterrò che l'interpretazione offerta da McGowan sia in ultima istanza da respingere poiché (i) presuppone un modello di sessualità assai problematico e (ii) espone la proposta dell'autrice al rischio di incoerenza.

Procediamo con ordine. Le richieste di permesso hanno comunemente a che fare con l'accesso a o l'uso di proprietà altrui, come in (3): per poter (essere in diritto di) usare il pc in oggetto A deve chiedere il permesso di B, che ne è il legittimo proprietario (e quindi la persona appropriata per accettare o rifiutare l'appello). A prima vista, queste considerazioni aderiscono senza grinze al caso delle avances sessuali: colui che approcci un'altra persona con l'intento di cominciare un rapporto sessuale non è in diritto di procedere oltre a meno che l'altro lo autorizzi a farlo, essendo il corpo di ciascuno sotto il proprio esclusivo controllo. L'analogia, però, non va molto lontano: nel caso del pc, A chiede a B il permesso per fare qualcosa che non coinvolge B direttamente (una volta ottenuto il permesso, A userà il pc *per conto suo*, per così dire); nel caso delle avances sessuali, invece, uno dei due chiede all'altro di fare qualcosa *insieme* (se l'altro acconsente, i due saranno impegnati in un'attività che coinvolge entrambi). Ritenere che, nel fare delle avances sessuali, uno avanzi delle richieste di permesso significa presupporre una visione delle relazioni sessuali in termini di *uso* da parte dell'uno del corpo dell'altro. Tipicamente, da parte dell'uomo del corpo della donna. In questa prospettiva, che fa da sfondo all'analisi di McGowan, il sesso è un'attività "a senso unico": è l'uomo a fare sesso con la donna, non viceversa – mentre le donne sono implicitamente ritratte come creature passive, prive di *sexual agency*.

Per lasciarsi, finalmente, tale modello alle spalle occorre reinterpretare le avances sessuali come tipi diversi di appelli aperti. Un buon candidato è l'atto di *proporre*: mentre le richieste

costituiscono tentativi di far fare qualcosa a qualcuno, le proposte rappresentano tentativi di coinvolgere l'altro in un'attività condivisa (cfr. Walton 2006). Occorre precisare che la presenza dell'elemento collaborativo, che pure è distintivo delle proposte, non ci garantisce che l'atto sia solo e tutto una proposta. Poniamo che *A* sappia che *B* ha appena ricevuto in regalo un enorme puzzle il cui completamento richiede parecchie ore di lavoro. *A* chiede a *B*:

(4) Lo componiamo assieme?

Mediante il proferimento di (4), *A* compie insieme una richiesta di permesso e una proposta. Chiede a *B* di poter avere accesso a una sua proprietà (richiesta di permesso) e, al contempo, lo invita a fare qualcosa insieme (proposta). Si noti che, nel caso in cui il puzzle fosse stato di proprietà di *A* (e non di *B*), lo stesso enunciato avrebbe costituito una semplice proposta (e non anche una richiesta di permesso). Ora, uno potrebbe interpretare le avances sessuali come proposte che implicano richieste di permesso, alla stregua di (4). Questa lettura, tuttavia, presuppone che nei contesti sessuali uno dei partecipanti chieda all'altro di poter "usare" il suo corpo (così come *A*, nell'esempio del puzzle, chiede a *B* di poterne usare le tessere), se pur al fine di svolgere un'attività congiunta. Lo snodo problematico è, ancora una volta, la riduzione del corpo – tipicamente, del corpo delle donne – a una proprietà di cui disporre come di qualsiasi altro possedimento (si veda in merito Du Toit 2009, pp. 35 e sgg.)<sup>16</sup>.

Alle proposte semplici e a qualsiasi tipo di appello aperto che non sia (anche) una richiesta di permesso seguono, lo dicevamo, rifiuti *non* autoritativi. *Contra* McGowan, per accettare o rifiutare avances sessuali, reinterpretate in termini di proposte semplici (o appelli simili), non si richiede l'autorità del parlante: Cindy può rifiutare le avances di Carl non tanto perché gode di autorità sul proprio corpo, quanto perché è la destinataria dell'appello. Cindy soddisfa un semplice requisito di pertinenza che Sally, al contrario, non rispetta. Se il rifiuto di avances sessuali non è un atto autoritativo, allora la riduzione al silenzio delle donne non può passare, come vorrebbe McGowan, attraverso la negazione di autorità pratica<sup>17</sup>.

L'implicazione di un modello di sessualità in aperto contrasto con gli scopi della causa femminista è un problema di natura innanzitutto etico-politica, ma la proposta di McGowan incorre anche in difficoltà di natura teorica. La nozione di riduzione al silenzio introdotta dall'autrice non sembra infatti coerentemente applicabile al rifiuto di avances sessuali. Per capire perché, ammettiamo come valida, per solo amor di argomento, l'analogia tra il caso del pc e il caso delle avances. Si noti che la forma di una richiesta di permesso tipo (si prenda (3) a mo' di esempio – "Posso usare il tuo pc?") è già di per sé indicativa del fatto che chi chiede sa di poter procedere solo previo consenso altrui. Una richiesta di permesso che sia genuinamente tale presuppone per definizione il riconoscimento, da parte del parlante, dell'autorità dell'altro sul dominio rilevante. Nell'avanzare una richiesta di permesso, il parlante assume che l'attività

<sup>16</sup> Ringrazio un revisore anonimo per aver elaborato una variante dell'esempio del puzzle e avermi esortata a specificare che certe proposte implicano richieste di permesso e pertanto richiedono che chi le accetta o rifiuta goda di autorità sul dominio rilevante.

<sup>17</sup> Benché si tratti di una questione largamente empirica non risolvibile facendo "filosofia in poltrona", ritengo plausibile che certe avances sessuali costituiscano richieste di permesso poiché così intese e così recepite. Al di là di che tipo di atto queste o quelle avances siano di fatto, la mia critica investe il modello teorico che fa da sfondo a quasi tutta la letteratura sul *silencing*: un modello incentrato sull'idea che un rapporto sessuale tipo cominci con una richiesta di permesso. Tale modello, come ho cercato di chiarire, implica una visione gerarchica della sessualità tanto ingiusta quanto erronea, secondo cui una parte (tipicamente, la parte maschile) desidera e agisce sulla base di ciò che desidera, mentre l'altra (tipicamente, la parte femminile) è passiva, dotata di agenzialità solo nella misura in cui questa è volta a compiacere il desiderio del partner. Per una critica simile ma rivolta alla letteratura su stupro ed etica sessuale, si veda Gardner (2017).

in questione ricada sotto la giurisdizione dell'interlocutore. Perché mai chiedere il permesso per qualcosa che uno potrebbe fare a prescindere dal via libera dell'altro? Che senso avrebbe cercare di ottenere il tuo consenso per poter usare ciò che è mio o che so non essere tuo? La mia richiesta, in circostanze simili, sarebbe quanto meno difettosa. Nel chiedere a B il permesso per usare un certo pc, A rivela allora la propria consapevolezza circa l'obbligo che ha di chiedere – ovvero riconosce che B, a differenza sua, può disporre del pc ed eventualmente cederlo a terzi. Se l'analogia tra i due casi è valida (e le avances costituiscono richieste di permesso, alla stregua di (3)), un uomo che faccia delle avances sessuali a una donna le riconosce *con ciò stesso* autorità sul proprio corpo. Riconosce il proprio obbligo di chiedere poiché lei (e solo lei) dispone del proprio corpo<sup>18</sup>.

Alla luce di quanto detto, si consideri nuovamente la tesi di riduzione al silenzio di McGowan: i “no” che alcune donne proferiscono in contesti sessuali sono votati al fallimento illocutorio poiché certi uomini, pur recependoli come tentativi di rifiuto, non riconoscono alle donne autorità sul proprio corpo. Ma se l'uomo, nel fare delle avances, sta “chiedendo il permesso” (come McGowan ritiene), l'ipotesi non regge: un uomo che chieda il permesso non può non riconoscere l'autorità del proprio interlocutore sul dominio in oggetto.

A partire dalla formulazione della celebre tesi anti-pornografia di MacKinnon, la nozione di *riduzione al silenzio* è divenuta materia di dibattito in filosofia del linguaggio femminista, ma anche in etica e, più recentemente, in epistemologia (cfr. soprattutto Fricker (2007) sulle diverse forme di ingiustizia epistemica, alcune delle quali si sovrappongono alle tipologie di riduzione al silenzio discusse sopra). Benché il rifiuto di avances sessuali sia stato posto al centro del dibattito come esempio paradigmatico di riduzione al silenzio delle donne, quasi nulla è stato detto in merito a cosa significhi rifiutare. In questo lavoro, ho cercato di offrire un'analisi del rifiuto in termini di atti linguistici. Dall'analisi è emerso che il rifiuto è un atto linguistico di “secondo turno”, ovvero un atto eseguito in risposta a una proposta, un invito, una richiesta, o altri tipi di appello “aperto”. Il rifiuto è un atto complesso, dal carattere ibrido, le cui caratteristiche e condizioni di felicità variano col variare dell'appello che lo precede. In particolare, ho sostenuto che il rifiuto è un atto autoritativo solo se preceduto da una richiesta di permesso. Alla luce dell'analisi condotta, ho poi valutato la proposta teorica di McGowan circa la riduzione al silenzio delle donne (cui contribuirebbero certi tipi di pornografia) – fenomeno che, a suo dire, passa attraverso il mancato riconoscimento, da parte di alcuni uomini, dell'autorità delle donne sul proprio corpo. Ho mostrato che concepire le avances sessuali maschili come richieste di permesso – e quindi le risposte alle avances come atti autoritativi – implica una visione asimmetrica delle relazioni sessuali e un'ingiusta (oltretutto falsa) rappresentazione delle donne come meri oggetti del desiderio maschile anziché soggetti dotati, al pari degli uomini, di *sexual agency*. La proposta di McGowan espone il fianco, inoltre, a critiche relative alla sua coerenza interna, dal momento che, se le avances sessuali sono richieste di permesso (come McGowan suggerisce), facendo un'avance si dimostrerebbe *ipso facto* di riconoscere l'autorità dell'altro sul proprio corpo – il che è esattamente ciò che McGowan sostiene non si dia nei casi di silencing.

Il dibattito sulla riduzione al silenzio delle donne mostra come la teoria degli atti linguistici di Austin possa essere applicata in maniere inedite e, fino a qualche anno fa, insospettabili.

## 5. Conclusione

---

<sup>18</sup> Ciò non vuol dire che l'uomo non possa in seguito ignorare il “no” della donna. Così come A, dopo aver ricevuto un secco “no”, può comunque usare il pc ignorando l'obbligo di non farlo impostogli dal rifiuto di B, allo stesso modo un uomo può ignorare (e, drammaticamente, si dà spesso il caso che gli uomini ignorino) il rifiuto di una donna, determinandone il fallimento perlocutorio.

Gli strumenti teorici messi a punto da Austin ci danno modo di indagare forme di ingiustizia passate largamente inosservate, che hanno a che fare con la sistematica limitazione del potenziale illocutorio di certi parlanti. A causa della propria identità di genere, etnica, religiosa o del proprio orientamento sessuale, certi individui possono fare con le parole meno e peggio di come sarebbero altrimenti stati in grado di fare. Tra le future direzioni di ricerca vi è, innanzitutto, l'indagine circa le strategie per fronteggiare tali forme di ingiustizia discorsiva (sulla nozione di ingiustizia discorsiva, si veda Kukla (2014)); per la delineazione di una possibile contro-strategia si veda Langton (in corso di stampa). Occorre inoltre elaborare un nuovo modello di sessualità, alternativo al cosiddetto "modello del consenso" (per una critica al modello del consenso si veda Anderson 2005). Indagine, questa, che può ancora una volta servirsi del "metodo degli atti linguistici", con il quale reinterprete le avances sessuali come atti altri rispetto a richieste di permesso, oltreché analizzare i numerosi atti linguistici coinvolti nei contesti sessuali, che non si limitano ad avances e risposte alle avances, ma includono svariati atti di negoziazione di ciò che i partecipanti intendono fare e come.

#### BIBLIOGRAFIA

- Anderson, M.J. (2005). Negotiating Sex, *Southern California Law Review*, 41, pp. 101-140;
- Austin, J.L. (1962). *How to Do Things with Words*, Oxford University Press, Londra;
- Bianchi, C. (2009). La parola, in N. Vassallo (a cura di), *Donna m'apparve*, Codice Edizioni, Torino, pp. 83-102;
- Butler, J. (1990). *Gender Trouble: Feminism and the Subversion of Identity*, Routledge, New York-Londra;
- Butler, J. (1997). *Excitable Speech: A Politics of the Performative*, Routledge, New York-Londra;
- Caponetto, L. (2016). Silencing Speech with Pornography, *Phenomenology and Mind*, 11, pp. 182-191;
- Coy, M., Kelly, L., Elvines, F., Garner, M., Kanyeredzi, A. (2013). 'Sex without consent, I suppose that is rape': How Young People in England Understand Sexual Consent, Office of the Children's Commissioner, London;
- Cowart, M. (2004). Understanding acts of consent: Using speech act theory to help resolve moral dilemmas and legal disputes, *Law and Philosophy*, 23, pp. 495-525;
- Du Toit, L. (2009). *A Philosophical Investigation of Rape. Making and Unmaking the Feminine Self*, Routledge, New York;
- Dworkin, R. (1991). Liberty and Pornography, *The New York Review of Books* (15 August 1991), pp. 12-15;
- Fricke, M. (2007). *Epistemic Injustice. Power and the Ethics of Knowing*, Oxford University Press, Oxford;
- Gardner, J. (2017). The Opposite of Rape, *Oxford Journal of Legal Studies*, DOI: 10.1093/ojls/gqx022;
- Green, M. (2015). Speech Acts: An Annotated Bibliography, *Oxford Bibliographies Online*, DOI: 10.1093/OBO/9780195396577-0300;
- Hornsby, J. (1993). Speech Acts and Pornography, *Women's Philosophy Review*, 10, pp. 38-45;
- Hornsby, J. (2014). Pornography and 'Speech', in L. Coleman, J.M. Held (a cura di), *The Philosophy of Pornography*, Rowman & Littlefield, Londra, pp. 129-146;
- Hornsby, J., Langton, R. (1998). Free Speech and Illocution, *Legal Theory*, 4, pp. 21-37;
- Kukla, R. (2014). Performative Force, Convention, and Discursive Injustice, *Hypatia*, 29(2), pp. 440-457;
- Lance, M., Kukla, R. (2013). Leave the gun; Take the cannoli! The pragmatic topography of second-person calls, *Ethics*, 123, pp. 456-478;
- Langton, R. (1993). Speech Acts and Unspeakable Acts, *Philosophy & Public Affairs*, 22(4), pp. 292-330;

- Langton, R. (2009). *Sexual Solipsism: Philosophical Essays on Pornography and Objectification*, Oxford University Press, Oxford;
- Langton, R. (in corso di stampa). Blocking as Counter-Speech, in D. Harris, D. Fogal, & M. Moss (a cura di), *New Work on Speech Acts*, Oxford University Press, New York;
- Levinson, S.C. (1983). *Pragmatics*, Cambridge University Press, New York;
- MacKinnon, C. (1987). *Feminism Unmodified: Discourses on Life and Law*, Harvard University Press, Cambridge;
- MacKinnon, C. (1993). *Only Words*, Harvard University Press, Cambridge;
- Maitra, I. (2009). Silencing Speech, *Canadian Journal of Philosophy*, 39(2), pp. 309-338;
- Malamuth, N.M., Hald, G.M., Koss, M. (2011). Pornography, Individual Differences in Risk and Men's Acceptance of Violence Against Women in a Representative Sample, *Sex Roles*, 66, pp. 427-439;
- McGowan, M.K. (2009). On Silencing and Sexual Refusal, *The Journal of Political Philosophy*, 17(4), pp. 487-494;
- McGowan, M.K. (2014). Sincerity Silencing, *Hypatia*, 29(2), pp. 458-473;
- McGowan, M.K. (2017). On Multiple Types of Silencing, in M. Mikkola (a cura di), *Beyond Speech: Pornography and Analytic Feminist Philosophy*, Oxford University Press, Oxford, pp. 39-58;
- Milburn, M.A., Mather, R., Conrad, S.D. (2000). The Effects of Viewing R-rated Movie Scenes That Objectify Women on Perceptions of Date Rape, *Sex Roles*, 43(9/10), pp. 645-664;
- Salih, S. (2002). *Judith Butler*, Routledge, New York-Londra;
- Saul, J., Diaz-Leon, E. (2017). Feminist Philosophy of Language, in E.N. Zalta (a cura di), *The Stanford Encyclopedia of Philosophy*, <https://plato.stanford.edu/archives/fall2017/entries/feminism-language/>;
- Sbisà, M. (2009). Illocution and Silencing, in B. Fraser, K. Turner (a cura di), *Language in Life, and a Life in Language*, Emerald, Bradford, pp. 351-357;
- Schegloff, E.A., Sacks, H. (1973). Opening up closings, *Semiotica*, 8(4), pp. 289-327;
- Searle, J.R. (1975). A Taxonomy of Illocutionary Acts, in K. Gunderson (a cura di), *Minnesota Studies in the Philosophy of Science*, VII, University of Minnesota Press, Minneapolis, pp. 344-369;
- Searle, J.R., Vanderveken, D. (1985). *Foundations of Illocutionary Logic*, Cambridge University Press, New York;
- Walton, D. (2006). How to Make and Defend a Proposal in a Deliberation Dialogue, *Artificial Intelligence and Law*, 14, pp. 177-239;
- West, C. (2003). The Free Speech Argument Against Pornography, *Canadian Journal of Philosophy*, 33(3), pp. 391-422;
- Wieland, N. (2007). Linguistic Authority and Convention in a Speech Act Analysis of Pornography, *Australasian Journal of Philosophy*, 85(3), pp. 435-456.





SECTION

5

# SECTION 5

HISTORY OF IDEAS

*Ethel Menezes Rocha*

God, Eternal Truths and the Rationality of the World in Descartes

*Marcos André Gleizer*

Spinoza on metaphysical doubt and the “Cartesian circle”

*Alfredo Gatto*

Descartes e il problema della teodicea nella prima modernità

*Raffaele Ariano*

Geografia e filosofia. Riflessioni su Pensiero vivente di Roberto Esposito a partire da Spinoza, Cavell e Foucault

*Corrado Claverini*

Dove va la filosofia italiana? Riflessioni sull'*Italian Thought*

---

ETHEL MENEZES ROCHA

Universidade Federal do Rio de Janeiro

ethel.rocha55@gmail.com

---

# GOD, ETERNAL TRUTHS AND THE RATIONALITY OF THE WORLD IN DESCARTES\*

---

## abstract

*In this paper I examine Descartes's thesis of the free creation of eternal truths in conjunction with what he claims to be the divine attributes known by us. Considering the Cartesian claims of God's simplicity and that eternal truths freely created by God include logical principles as the structure of finite minds, I argue that the Cartesian thesis of the free creation of eternal truths involves: a) God necessarily establishes (and thus creates) within/as himself all essences and truths, including those that to the finite intellect seem impossible; b) While establishing essences and truths, God instantiates (and thus creates) some of what to the finite intellect seems to be non-contradictory essences and truths as innate ideas in actual finite minds, and instantiates at least some of what seems to be non-contradictory essences and truths in the actual physical world and c) While establishing essences and truths, God may have instantiated in the actual world many (or all) of the essences and truths that the finite mind perceives to be contradictory, that is, those whose existence cannot be clearly and distinctly conceived by the finite mind and seem, therefore, impossible.*

---

## keywords

*God, eternal truth, world, human knowledge*

\* An initial version of this work was presented in 2013 at the Philosophy of Language Seminar – Federal University of Rio de Janeiro, UFRJ, Brazil; in 2014 I presented a second version in the research group U of T Modern Philosophy Research Group - University of Toronto, UofT, Canada. Still a third version was presented at the History Workshop at UCLA, USA. The rewording of the original text from the inciting discussions and also the precise comments resulted in this text. Throughout these periods, I obtained support from CNPq and from CAPES, to develop my research at the Federal University of Rio de Janeiro, Brazil and at the University of Toronto, Canada.

**1. Introduction** It is well known that Descartes advances the strong thesis that God freely creates eternal truths. With this thesis, Descartes does not merely argue that everything (including eternal truths) depends on God, as is sustained by the Aristotelian scholastic tradition, but rather that everything (including eternal truths) depends on God's free will, such that God could have created things differently. This proposition does not appear in the text of any of Descartes's published books, but it does emerge in his correspondence from 1630 and 1649 and in his replies to two of the objections made to the *Meditations on First Philosophy*. It also appears in *Conversation with Burman*, which consists of Frans Burman's notes of his interview with Descartes that took place in 1648.

In this paper, I examine Descartes's thesis of the free creation of eternal truths in conjunction with what he claims to be the divine attributes known by us. I focus on the meaning of Descartes's claim that God "could have created things differently", considering that the eternal truths freely created by God include the logical principle of non contradiction, that is, the basic principle of human rationality. I argue that *as far as finite human intellects can conceive*, the Cartesian thesis of the free creation of eternal truths involves at least the following: a) God, being infinite and pure thought in act, *necessarily* establishes (and thus creates) within/as himself all essences and truths, including those that to the finite intellect seem impossible, since they appear contradictory; b) God, being omnipotent and/or indifferent, while establishing essences and truths, instantiates (and thus creates) some of what to the finite intellect seems to be non-contradictory essences and truths as innate ideas in actual finite minds, minds which are created with logical structures, and also instantiates in the actual physical world at least some of these seemingly non-contradictory essences and truths instantiated in finite minds; and c) although the finite human mind cannot conceive of them clearly and distinctly, God, while establishing essences and truths, may have instantiated in the actual world all of the essences and truths that the finite mind perceives to be contradictory, that is, those whose coexistence cannot be clearly and distinctly conceived by the finite mind and seem, therefore, impossible.<sup>1</sup> I will then turn to two examples in Descartes writings which confirm that Descartes is willing to admit that God has created in the world

---

<sup>1</sup> For Descartes's view that when a finite mind clearly and distinctly perceives an idea, it at least understands the meaning exhibited by the idea's content, and that this means that it perceives this content as not contradictory and that what is exhibited can exist in the world see, for example, AT 6: 71; CSM 2: 50.

things that the limited, finite mind understands as impossible to exist. Thus, my main claim here is that by introducing the free creation doctrine, Descartes seems to have allowed for an absurd actual world. This doctrine opens the possibility that God has created a world that cannot be understood by human basic rational principle.<sup>2</sup> Further, this is so not with respect to the mysterious truths of faith, such as transubstantiation or the Holy Trinity, which, according to Descartes, in spite of being obscure and confusing, can be clearly and distinctly accessed<sup>3</sup> through divine illumination. My claim here is, rather, that the free creation doctrine opens the possibility of created truths (that is, truths from God's point of view and, therefore, truths in themselves) which, if grasped by human beings, are grasped through obscure and confused ideas in such a way that they can only appear as contradictory and absurd. Through the free creation doctrine Descartes admits the possibility of God's creating truths about which human beings cannot have clear and distinct ideas, either if regarded by reason. Moreover, based on the fact that, according to Descartes, we do have obscure and confused access at least to some of what seems to us to be the contradictions that are created by God, I contend that Descartes admits that God indeed created a world in which at least some of what human beings perceive as contradictions are truths. Although Descartes does not expressly argue for the creation of a meaningless world, the fact that he does admit that God creates at least some truths that seem absurd to finite minds confirms that the created world may contain many truths (since it contains at least some) that, according to him, the finite mind sees as contradictory or that are "beyond its understanding."

The paradigm of the eternal truths that are indicated by Descartes in different texts is mathematical truths. However, Descartes also includes in the list of eternal truths logical principles, as well as metaphysical, physical, and moral principles. The inclusion of logical principles among the things freely created by God is clear in various passages, such as, for example, the following from the *Principles of Philosophy*:

[T]he proposition, *Nothing comes from nothing*, is regarded not as a really existing thing, or even as a mode of a thing, but as an eternal truth which resides within our mind. ... The following are examples of this class: *It is impossible for the same thing to be and not to be at the same time; What is done cannot be undone; He who thinks cannot but exist while he thinks*; and countless others. [AT 8A: 24; CSM 1: 209]<sup>4</sup>

Along with the claim that the cited examples are eternal truths, this passage suggests that logical principles are created as the structure of the finite intellect, since Descartes states that, among others, the proposition "*It is impossible for the same thing to be and not to be at the same time*" must be considered as referring not to *really existent things*, nor as a *mode of something*, but rather as a certain eternal truth that has its locus *in our minds*. Thus, logical principles are created in the finite mind, though not as a contingent modification of it: it is not as if the finite

## 2. Eternal Truth, Logical Principles, and the Finite Intellect

<sup>2</sup> For opposite views refer to Ishiguro (1986) and Normore (1991).

<sup>3</sup> For Descartes's discussion on the clarity and distinction of the ideas that come from grace, see his letter to Dinet [AT 7: 581; CSM 2: 392], his "Reply to the Second Objections" [AT 7: 147/148; CSM 2: 105], and his letter to Hyperaspistes [AT 3: 426; CSMK: 191].

<sup>4</sup> References to Descartes' writings indicate the edition of *Oeuvres de Descartes* ed. by C. Adam & P. Tannery (AT), followed by the translation of Descartes works by J. Cottingham, R. Stoothoff, and D. Murdoch, *The Philosophical Writings of Descartes* (CSM), both followed by their respective volume and page numbers, or the translation of Descartes's correspondence by J. Cottingham, R. Stoothoff, D. Murdoch, and A. Kenny in *The Philosophical Writings of Descartes* (CSMK), followed by page numbers.

mind is created in such a way that it sometimes thinks logically, as it sometimes doubts, loves, hates, and so on. In a letter to Arnauld, Descartes clarifies it: the logical principles configure the way the finite mind can think, that is, they constitute the finite mind's structure. As he writes:

I would not dare to say that God cannot make a mountain without a valley, or bring it about that 1 and 2 are not 3. I merely say that he *has given me such a mind that I cannot conceive* a mountain without a valley, or a sum of 1 and 2 which is not 3; such things involve a contradiction *in my conception*. [AT 5: 224; CSMK: 358/359; emphasis added]

Moreover, as Descartes argues in the Fourth Meditation, this logical structure of the finite mind is what makes it limited. This becomes clear when Descartes asserts that the finite mind is endowed with two faculties: understanding, which is finite, passive, and responsible for conceiving ideas, and will, which is infinite, active, and responsible for every action, including any mental act that adds to the representations of understanding, such as when one wishes for something, asserts something, doubts something, and so on. As Descartes says, it is the infinitude of the will that makes the finite mind similar to God and the finitude of understanding that limits it [AT 7: 57-8; CSM 2: 39-40].

Since logical principles are created as the structure of the finite mind and as what limits it, it is reasonable to assert that the very concept of contradictoriness may be related solely to the human mind.<sup>5</sup> What appears contradictory to the finite mind, and thus impossible to exist, seems to result from a limitation of such a mind that does not necessarily correspond to how things are in fact, that is, how things are from the viewpoint of the creator. Things created by God and, thus, from God's own viewpoint, can always coexist with each other. Thus, there exists a way that things are in fact, that is, as they are created by God, and in this way everything may coexist with everything else. The human mind, nevertheless, cannot perceive them as such. On the one hand, contradictions seem to result from the obscure way in which the finite mind combines certain ideas, as indicated by Descartes to Burman:

[T]here is no contradiction in things, but in our ideas alone. For it is ideas alone that we join together *in such a way* that they are inconsistent one with another. Things, by contrast, are not inconsistent with each other, since all of them can exist: so no one thing is inconsistent with any other. With ideas, the opposite is the case: in our ideas we join together and unite separate things, which taken on their own are not inconsistent. This is the origin of contradiction. . . . [T]he *conception you have of the combination and unity of the two ideas is not clear but extremely obscure*. (AT 5: 38, emphasis added)<sup>6</sup>

It seems that Descartes here is not only saying that things are distinct from each other and, therefore, in themselves, as God creates them, are not inconsistent with each other, but also that the inconsistency that the finite mind perceives concerning the coexistence of some of them results from its limited and muddled way of thinking. Thus, it is the finite mind that sees the coexistence of certain things as impossible, and this is not because it combines what cannot be combined, but because it combines ideas about things in a very obscure way. From God's point of view nothing is such that it cannot be, which seems to mean that it is

---

<sup>5</sup> This line of interpretation is suggested by Wilson (2005), p. 110.

<sup>6</sup> *Descartes' Conversation with Burman*, trans. Cottingham (1976), p. 25. This passage was not included in the CSMK volume.

possible that God can make anything coexist with anything. For example, in a letter to More, Descartes writes that he cannot say that God cannot do what he sees as contradictory:

And so I boldly assert that God can do everything which I perceive to be possible, *but I am not so bold as to assert the converse, namely that he cannot do what conflicts with my conception of things – I merely say that it involves a contradiction.* [AT 5: 272; CSMK: 363]

Similarly, in a letter to Mersenne, he declares, “In general we can assert that God can do everything that is within our grasp *but not that he cannot do what is beyond our grasp*” [AT 1: 146; CSMK: 23 ]; again, in a letter to Mesland, he says, “God cannot have been determined to make it true that contradictories cannot be true together, and therefore . . . he could have done the opposite” [AT 4: 118; CSMK: 235].

If God created these logical principles as the structure of finite intellects that limit them, then it is at least possible that in the act of creation God is not himself subject to logical principles, and this implies that God may indeed have created truths that the finite mind cannot conceive.

Up to now, as we have seen, it can only be asserted that God, in creating, may or may not be subject to created logical principles from all eternity. God creates logical principles in the finite mind but, it would seem, nothing stops him from following them in his single act of creation of everything. God could create logical principles and simultaneously use them throughout all his creation. It seems that given his omnipotence nothing stops him and, likewise, nothing obliges him to use created principles while creating everything, including these principles. However if we focus on what Descartes says we know of God’s nature, we are led to argue that he must admit the stronger thesis, according to which there is a sense in which God *cannot* be subject to logical principles. This is so not because God is sufficiently powerful not to be subject to what he creates. His omnipotence allows that he creates from eternity and is subjected to what he creates. Rather, we must admit, God cannot be subjected to the logical principles for two reasons that concern what, according to Descartes, we know about God’s very infinite essence. Note that, in spite of sustaining the incomprehensibility of God’s infinite power by the finite mind, Descartes also sustains that this finite mind can know at least some of God’s attributes:

But there are many things really in God, or which relate to God, which we are capable of reaching in our minds and expressing in word . . . And so in this sense God can be known and spoken of to a very great extent ... [AT 2: 284, CSMK: 169]

Before examining these two reasons, it is worth saying a word about what this overall reading of Descartes suggests. Since the Cartesian God is the creator of all things, including the finite mind’s structure and its particular thoughts (in the sense that he creates our mind with the disposition to think whatever particular thoughts we think), it is clear that anything a finite mind thinks that it knows about God, essences, minds and the actual world is governed by the operator “I clearly and distinctly perceive that.” What we think we know about anything is what God allows us to know. Further, this is so in such a way that, strictly speaking, it is the limitation of a finite mind that makes it consider that God cannot choose not to be subjected to logical principles. That is the way the finite mind perceives it. It is worth to mention that, according to Descartes, there is no way a finite mind can access what would be an absolute truth or an absolute falsity. What is perceived as unquestionably true by the finite mind is that which it clearly and distinctly perceives, according to its logical structure. However, if this is so, our perceptions of truth are the best of which we are capable of perceiving clearly

### 3. God’s Nature and Logical Limitation



and distinctly. This seems to be clear, for example, in Descartes' "Reply to the Second Set of Objections" to the *Meditations*:

First of all, as soon as we think that we correctly perceive something, we are spontaneously convinced that it is true. Now if this conviction is so firm that it is impossible for us ever to have any reason for doubting what we are convinced of, then there are no further questions for us to ask; *we have everything we could reasonably want*. What is it to us that someone may make out that the perception whose truth we are so firmly convinced of may appear false to God or an angel, so that it is absolutely speaking, false? *Why should this alleged 'absolute falsity' bother us, since we neither believe in it nor have even the smallest suspicion of it?* [AT VII: 144; CSM II: 103, emphasis added]

The first reason by which we understand that God cannot be subject to logical principles is based on the fact that we know, as becomes clear in Descartes's first proof God's existence, that he is infinite, which means that he has no limit of any kind. The formal reality of an infinite being is derived from the fact that a finite mind cannot be the cause of the infinite objective reality depicted in the idea of God. Besides that, we also know that what is infinite has no kind of limitation. As Descartes says in his reply to Caterus's objection concerning the possibility of a clear and distinct idea of the infinite: "It [the infinite] can still be understood, in so far as we can clearly and distinctly understand that something is such that *no limitations can be found in it*, and this amounts to understanding clearly that this thing is infinite" [AT 7: 112; CSM 2: 81; italics mine]. Since what we clearly and distinctly understand about the essence is true about the thing that has this essence, as Descartes asserts in the Fifth Meditation [AT 7: 65; CSM II: 45], then, to understand that God is infinite means that God is infinite. And, since we know that God is infinite, we know that he has no limitations. The absence of any limits in God is also clearly affirmed in the *Principles*, when Descartes distinguishes what is indefinite, or anything in which *we do not perceive a limit*, from what is infinite, or God, in which *we perceive that there is no limit*: "For in the case of God alone, not only do we fail to recognize any limits in any respect, but our understanding positively tells us that there are none" [AT 8A: 15; CSM 1: 202]. Now, if God is infinite in the sense in which he has no limits and if a logical structure is a limitation (of the created finite mind), then it is not that God may or may not be subject to logical principles, but rather that he necessarily is not subject to them. Given his infinitude, God could not choose to be subject to them as he could not choose not to exist, not to be omnipotent, and so on, since he has no limits at all.<sup>7</sup>

The second reason that leads me to assert that, in the Cartesian system, God, in creating, cannot be subject to logical principles is also based on what, according to Descartes, we know about God's infinite nature: his simplicity, unity, immateriality, and actuality. In the Third Meditation Descartes writes: "[T]he unity, the simplicity, or the inseparability of all the attributes of God is one of the most important of the perfections which I understand him to have." [AT 7: 51; CSM 2: 34] Note that, as well as affirming the simplicity and unity of God, Descartes, in this passage, also clearly admits that he understands that God has different perfections (in the plural), albeit inseparably. In other words, Descartes does not seem to be upholding the thesis that *we cannot distinguish them* in God. While the thesis of the simplicity and unity of God leads to the claim that in God there is no plurality of properties, it also seems to allow that we can, in some way, distinguish different aspects of God's infinitude.

---

<sup>7</sup> For a discussion on the possibility of God choosing his nature, refer to Gueroult (1953) and Schmaltz (1991). Though based on different arguments, they both claim that God's nature cannot be altered even from eternity.

There is not a plurality of attributes in God, although we may refer to his single attribute – infinity--through different aspects of it (omnipotence, immutability, benevolence, etc.). Thus, Descartes appears to claim that while there is no real or modal distinction between the different properties *we perceive* in God, we distinguish them mentally, that is, conceptually. In *Principles* I, 60 [AT 8A: 29-30; CSM 1: 213], Descartes explains that a conceptual distinction, or distinction of reason, involves the basic idea that we can use different concepts to refer to essentially connected things or a single thing. If God is one, his properties are not separate in him. Although simple, God can be conceived of in diverse ways. Further, although we conceive of him in diverse ways, each of these ways is indeed it and all the others together. Because we know that God is infinite, we think of him as simple and one and as immutable, indivisible, immaterial, and pure actuality: He has no parts or modes and nothing in him is in potency. As Descartes puts it: “[G]od, who is the author of all things, is entirely perfect and unchangeable” [AT 3: 649, CSMK: 216/17]; “[S]ince being divisible is an imperfection, it is certain that God is not a body” [AT 8A: 13; CSM 1: 201]; “[G]od, on the other hand, I take to be actually infinite, so that nothing can be added to his perfection” [AT 7: 47; CSM 2: 32]. Thus, we know that God is a single, unchangeable, pure immaterial act, which, given Descartes’s economic ontology of two kinds of substances (the *res cogitans* or mind or soul and the *res extensa* or extended substance or body) and their modes, means that he is a single, pure intellectual act. As Descartes says, “God is pure intelligence.” [AT 10: 218; CSM 1: 5] God is, then, according to Descartes, the pure infinite thinking substance. Further, as a consequence of his simplicity, thinking, wishing, and creating are in him a single act, as Descartes asserts in, for example, the *Principles*: “[T]here is always a single identical and perfectly simple act by means of which he [God] simultaneously understands, wills, and accomplishes everything” [AT 8A: 14; CSM 1: 201]. God is at least a pure act of thinking/wishing/creating things. Hence, based on the attributes that we know of God, we can at least assert (1) that these attributes are identical to each other; (2) that God is infinite and pure thinking in act; and (3) that God’s action is eternal and singular.

Now, besides the attributes of God that Descartes admits to us knowing, we also recognize that they are in an infinite degree, so that, for example, what we understand as his understanding is an infinite understanding, as is clear in the Fourth Meditation:

If, for example, I consider the faculty of understanding I immediately recognize that in my case it is extremely slight and very finite, and I at once form the idea of an understanding which is much greater – indeed supremely great and infinite; and from the very fact that I can form the idea of it, I perceive that it belongs to the nature of God. [AT 7: 57; CSM 2: 39-40]

Because of God’s infinite understanding, it can be affirmed, then, that God thinks of all things. However, if thinking, willing, and creating are a single act then, by thinking all things, God wills and creates all things. The question then becomes what it means to say that God thinks all things. Now, in this context, “God thinks all things” could mean that God thinks of everything that the finite intellect, structured by logical principles, thinks and only this. If so, then we would have to admit that God is necessarily subject to logical principles, since what is thinkable for the finite intellect is what is in accordance with its finite constitution. Were God to think only like us, he would be, like us, subject to the finite constitution of our minds, which, as we have seen, would be against the thesis of God’s infinitude. However, “God thinks all things” could also mean that God thinks whatever he wishes, thereby limiting God’s thinking to his will. If this were so, we would have to admit that there is in God a faculty, namely the will, that is not only conceptually distinguished from the intellect, but is

paramount over it, which would contradict the thesis of the identity of his attributes. Hence, it seems, we have to admit that God “thinking all things” means that God’s thought is not limited in any way, which implies that he *necessarily* thinks what for us is thinkable and thus logically possible *and* what for us is unthinkable and thus logically impossible. Furthermore, if by thinking, God wills and creates, then there is a sense in which God *necessarily* thinks, wills, and creates what is thinkable by the finite intellect and what is not thinkable by it.

This reading of the Cartesian thesis of the free creation of eternal truth, according to which *there is a sense* in which God necessarily creates all the essences and truths, seems to collide with the very basis of the free creation thesis, that is, that God freely creates these essences and truths. Nonetheless, I suggest, based on the Cartesian concept of God’s nature, which involves his immateriality, infinite power, and indifference, it is reasonable to say that, this is only an apparent collision. According to Descartes, God’s freedom involves a positive indifference in such a way that he acts with no external determination. Descartes thus states, “It is self-contradictory to suppose that the will of God was not indifferent from eternity with respect to everything which has happened or will ever happen.” It is exactly because of this, that is, because God acts without any external determination from what was or will be the case that, as Descartes argues in the same passage, “the supreme indifference to be found in God is the supreme indication of his omnipotence” [AT7: 431/432; CSM2: 291]. Now, if God’s freedom consists at least in his acting without external coercion, to say that in following his own nature (of being a pure act of thinking all things), God necessarily thinks/creates all essences and truths means that God freely, though necessarily, creates these essences and truths. As Descartes explains to Burman, “We should not make a separation here between the necessity and the indifference that apply to God’s decrees: *although his actions were completely indifferent, they were also completely necessary*” [AT 5: 167; CSMK 347].

#### 4. God’s Nature and Creation

Thus, as we saw, given his infinitude, God has no limits in any aspect and what he thinks, he wills and creates. In what follows I argue that God creates essences and truths in himself, as his own thoughts, which means that he causes himself, and in the same single act, he instantiates some in finite minds and some (or all) in the actual world.

Now, if, as stated above, God is a single pure thinking in act, then his essence is constituted, so to speak, “of the contents of his thought.” However, if God is simple and immutable, he has no parts or modes and, therefore, his act of thinking and the contents exhibited by this act are one and the same thing. As we saw with respect to all the different attributes we think God to have, the difference between God’s act and God can be only a distinction of reason, that is, a distinction that does not correspond to a distinction within God, but only a way the finite mind thinks about God. God has no modification, so the different acts which we might think of as his different acts are indeed identical to what he is. More than that, if there are no modifications or division in God, then whatever is involved in God’s act of being, is also identical to God.

Further, as we saw, God is a thinking act, so he is identical to this act and to whatever is involved in it. There is no such a thing as a content exhibited in what would be God’s intellect when he thinks. God is identical to a single pure act of thinking which does not depict any content but, rather, constitutes his own essence. God is identical to his own thoughts, which, strictly speaking, from God’s point of view, consist in one single act/content of being thought from eternity.<sup>8</sup> Nevertheless, if, as we saw, what would be the “content of his thought” are all essences and truths, then God is the same as these essences and truths. This would appear to be confirmed by what Descartes says to Burman:

---

<sup>8</sup> For a close but different approach on the plausibility of this thesis see Schmalz (2009) and Rozemond (2008).

Whatever is in God is not in reality separate from God himself; rather it is identical with God himself... Then, again, although we may conceive that the decrees could have been separated from God, this is merely a token procedure of our own reasoning: the distinction thus introduced between God himself and his decrees is a mental one not a real one. In reality the decrees could not be separated from God: he is not prior to them or distinct from them, nor could he have existed without them. [AT 5: 166; CSMK: 348]

If this is so, that is, if due to his immutability, simplicity, unity, and immateriality, God's creating essences and truths as "his thoughts" means that he creates them as himself, the essence of everything, then we have to admit that God is self-caused, which is indeed one of Descartes's innovating thesis in relation to tradition. The fact that the two theses--that is, the free-creation and the *causa sui* theses--are in an important way connected suggests that the Cartesian thesis of God as *causa sui* is not simply a consequence of his more general approach to causation (as asserts, for example, Gilson and latter Machamer & McGuire<sup>9</sup>), but rather an exigency of his free creation thesis: God's self-causation thesis, seems to be the only way available to guarantee the thesis that God freely creates the eternal truths which, as we saw, are identical to himself. However, this discussion is beyond the scope of this paper. Furthermore, according to Descartes, God in a single act freely creates not only essences and eternal truths, but also all that exists. This is what he says, for example, in his "Reply to the Third set of Objections": "[I]n demonstrating the existence of God we have also demonstrated that God created the entire world, or all things which exist apart from him" [AT 7: 188; CSM 2: 132]. Or, as Descartes says in his "Reply to the Sixth Objections," "[T]here can be nothing whatsoever which does not depend on [God]. This applies not just to everything that subsists, but to all order, every law, and every reason for anything's being true or good... [T]hey [eternal truths] depend on God alone, who, as the supreme legislator, has ordained them from eternity" [ATVII: 435; CSM II: 293-4]. As we saw, given his omnipotence and simplicity, God is self-caused and is what he thinks, that is everything (albeit eminently). If God is the cause of everything while thinking, and if there are other things besides God, then along with creating his infinite being as he thinks, God also creates what is finite. Thus, while because of his simplicity, whatever God thinks he creates, it seems to be the case that God in the very same act, in eternity, creates different essences with different modes of existence: as himself, some instantiated in finite minds, and some (or all) instantiated in the actual world. If God, is pure infinite thought in act, he is essentially (albeit eminently) all the things thought by him. In strict sense, says Descartes, God *creates* existences (in finite minds and/or in the actual world) and *establishes* eternal essences in himself as himself:

You ask what God did in order to produce them [the eternal truths]. I reply that from all eternity he willed and understood them to be, and by that very fact he created them. Or, if you reserve the word "created" for the existence of things, then he established them and made them. [AT 1: 152-53; CSMK: 25]

As for created minds, since they are made as limited by logical principles, it is clear that God does not endow them with everything he thinks: "[I]t is not just that our understanding ranges over fewer objects than that of God: rather, it is extremely imperfect in itself, being obscure, mingled with ignorance, and so on" [AT 5: 158; CSMK: 341]. The finite mind is, therefore, as we saw, structurally limited. It not only cannot think certain ideas that God thinks, but

---

<sup>9</sup> Gilson (1930) and Machamer & McGuire (2009).

even among the ones it can think, some cannot be thought clearly and distinctly. However, unlike the case of finite minds, Descartes does not seem to claim that the extended world is structurally limited. It is certain that the extended world is finite for, according to him, God creates any existent substance apart from him, which means that just by being created any substance is not infinite. The extended substance, as any other created thing, is finite in this sense. But Descartes seems to believe that, unlikely the thinking substance, the extended substance is finite but not structurally limited. All he says is that since we cannot clearly and distinctly perceive its limit, it might or might not have limits:

Having then no argument to prove, and not even being able to conceive, that the world has bounds, I call it *indefinite*. But I cannot deny on that account that there may be some reasons which are known to God though incomprehensible to me; that is why I do not say outright that it is *infinite* [AT 5: 52; CSMK: 320]

Thus, according to Descartes, though we recognize the finitude of the world in the sense that it is created, we cannot find any structural limit in it, and this means that it might have no limit. In the same way that it would be a limitation of God's infinitude, if he could not think what to the finite and limited intellect is unthinkable, it would also be a limitation, if he had to create as existent in the actual world everything he thinks or if he could not create in the actual world everything he thinks. Thus, if we cannot find limits to the extended world and if there is no limit to God's power, then, at least as a hypothesis, God may have put into the actual created world all that he thinks. Nothing determines God to think into the world only what we think he does.

Therefore, on the one hand, in spite of the simplicity of God, according to which what God thinks, he wills and creates, thinking, willing, and creating eternal truths does not mean necessarily creating them in the actual world, nor in the finite mind. On the other hand, though, it is reasonable to say that God may have created in the world everything he thinks or, at least, something that the created mind cannot conceive. God thinks everything and therefore creates as himself every essence and truth, but depending on the way he thinks them, these essences and truths have (or have not) counterparts in the physical actual world or in the finite mind. Further, if the finite mind is limited and God is not, it is possible that not everything God creates is knowable by this mind. As Descartes says, "And how do we know that God has not produced an infinite number of kinds of creatures, and thus, as it were, poured forth his power in the creation of things?" [AT 5:168; CSMK: 349]

Now, if, on the one hand, Descartes does not expressly say that he thinks that God has indeed created in the actual world everything that he thinks but that is unconceivable to the finite mind, on the other hand, he mentions at least two cases that indicate that he believes that God has indeed created in it some of these things. According to Descartes, we clearly and distinctly perceive that God pre-ordains everything from eternity, and we also clearly and distinctly perceive the real distinction between the body and the soul. Besides that, we have intimate experiences of our freedom and of the union of our bodies to our minds. This means that somehow we have access at least to these pairs of what we clearly and distinctly perceived as contradictories:<sup>10</sup> a pre-ordained world where there are free beings and a being that is a union

---

<sup>10</sup> There is an enormous amount of secondary literature about the possibility of dissolving these, at least apparent, pairs of contradiction (for a list of it, refer to C.P. Ragland (2005). In this paper, however, I am taking seriously what, as we will see, Descartes expressly says in this connection: those theses (of our freedom, God's preordination of everything, real distinction between body and soul, and their intimate union) appear to us as contradictories and, therefore beyond our comprehension, but, in spite of that, correspond to what God created in the actual world.

of two things that exclude each other.

In a letter to Elizabeth of Bohemia, Descartes suggests that at least one of God's creations appears to the finite intellect as involving a contradiction, which therefore, cannot be clearly and distinctly conceived: the union of body and soul, the union of two real distinct and opposite substances. He says, "It does not seem to me that the human mind is capable of forming a very distinct conception of both the distinction between the soul and the body and their union; for to do this it is necessary to conceive them as a single thing and at the same time to conceive them as two things; and this is absurd" [AT 3: 693; CSMK: 227]. Although, according to Descartes, the finite mind cannot conceive the possibility, God has created human being as the union of body and soul, that is, the union of two really distinct substances that are "not only different, but in some way opposite" [AT 7: 13; CSM 2: 10].<sup>11</sup> As it is clear from the Sixth Meditation, the finite mind has a clear and distinct idea that body and soul are really distinct [AT 7: 78; CSM 2: 54] and exclusive of one another. However, because of sensible experiences such as pain, thirst, and so on, the finite mind perceives that it is closely joined to a particular body that forms with it a unit. Thus, says Descartes, the union of body and soul can "only obscurely [be conceived] by the intellect alone or even by the intellect aided by the imagination, but it is known very clearly by the senses" [AT 3: 691; CSMK: 227]. The way the finite mind has access to the union and to the distinction of body and soul is different but, nevertheless, they are both truths. Nevertheless, as shown by the passages mentioned above, though they are both true at the same time, the finite mind cannot make sense of or clearly and distinctly understand them being true at the same time. The contradiction rests in the fact that through our finite minds we know that mind and body are two distinct and exclusive substances and, in spite of that, through our bodies we apprehend that they are in an intimate single substantial union. The fact that the real distinction is known through a clear and distinct perception of the mind and that the union is known through confused sensations does not preclude that both are equally true and known. If this is so, human beings somehow have access to at least this pair of truths that, nevertheless, they perceive as a contradiction. Another indication that Descartes thinks that God has indeed created as existing in the actual world things that appear as contradictory to our finite mind or, as he says, things that are "beyond our comprehension" is found in the *Principles*, when in introducing the question of the possible compatibility between the theses of divine preordination and human liberty, Descartes asserts, "[W]e can easily get ourselves into great difficulties if we attempt to reconcile this divine preordination with the freedom of our will, or attempt to grasp both these truths at once" [AT 8A: 20; CSM 1: 206]. As for the preordination, writes Descartes,

[A]ll the reasons that prove that God exists and is the first and immutable cause of all effects that do not depend on human free will prove similarly, I think, that *he is also the cause of all effects that do so depend...* [P]hilosophy by itself is able to discover that the slightest thought could not enter into a person's mind without God's willing, and have willed from all eternity, that it should so enter. [AT 4:314; CSMK: 272, emphasis added]

Therefore, through a clear and distinct idea, the finite mind knows that God exists and is the creator of all the finite mind's thoughts, which includes its will. In spite of this, Descartes also holds that "the supreme perfection of man is that he acts freely or voluntarily, and it is this which makes him deserve praise or blame" [AT 8A: 18; CSM 1: 205]. Moreover, the finite

---

<sup>11</sup> It is clear from Descartes's identification of the substance with its principal attribute that he admits body and soul to be exclusive of one another. Refer to AT 8B: 349; CSM 1: 298.

mind also has a clear and distinct idea of this freedom, which is self-evident, as he says in the *Principles*: “[W]e have such close awareness of the freedom and indifference which is in us, that there is nothing we can grasp more evidently or perfectly” [AT 8A: 20; CSM 1: 206]. Thus, according to Descartes, God is so powerful that somehow he preordains free choices, though we cannot conceive how this is so, since their coexistence seem to be contradictory and, therefore, impossible on our viewpoint. But, says Descartes, “[I]t would be absurd, simply because we do not grasp one thing, which we know must by its very nature be beyond our comprehension, to doubt something else of which we have an intimate grasp and which we experience within ourselves” [AT8A: 20; CSM 1: 206]. Again, although the finite mind cannot clearly and distinctly conceive that at the same time both that human being is free and that God preordain everything, that is how God has created the world.<sup>12</sup>

Thus, based on his free creation and God’s infinite nature doctrines, it is reasonable to think that Descartes admits that in a single and very complex act of being/thinking God creates essences and eternal truths in three different modes of existence: as himself, in the actual world, and in actual finite mind (and they do not necessarily match each other). If this is so, we may conclude that it is also reasonable to say that, for Descartes, God has actually created in the world at least some (and might have created all) of the things that he thinks but that seems impossible to the finite mind.

**5. Final Considerations**

We have concluded that, according to Descartes, in a single pure act of thinking, God exists and thus establishes essences as his infinite intellect (which is the whole of himself), instantiating in the finite intellect some of them (those that are logically conceivable) and instantiating in the formal world, some or all of these essences thought and established by and as himself. The finite intellect can clearly and distinctly perceive that God necessarily creates all the essences as himself, instances some of these essences as the content of ideas in finite minds, instances some of them in the actual world, but perhaps it cannot clearly and distinctly perceive all the essences God instantiates in the world.

The interpretative hypothesis of the thesis of the free creation of the eternal truths presented here seems to lead at least to the following problem: If it is correct, then we have to sustain that Descartes admits the possibility of an actual world created by God (or part of it) that makes no sense to the finite mind, that is, a world unintelligible from the viewpoint of human beings. Given this hypothesis that God necessarily thinks of all essences and truths and also given that God may have thought them all as actually existing, it is, perhaps, necessary to admit that in the Cartesian system God, is deceptive, in a certain sense: he structures the human intellect with cognitive faculties that operate according to logical principles, which, however, might prevent this intellect from knowing the actual world. God would not be deceptive in the strict sense of implanting clear and distinct false ideas in the finite mind, but deceptive in another sense, namely, that despite creating the human mind with principles and a conceptual system to know allegedly the world, which seems to imply that the world is rationally comprehensible, God creates things in the world that the finite mind, structured by him in this way, is incapable of knowing.

If God freely creates the logical principles in the finite minds and is not himself subject to it in his only act of creating everything, then it is plausible to admit that in spite of the fact that finite minds can not conceive how, it might be the case in the actual world not only that what is clear and distinct known is true, but also that its contradiction is true in God’s point of view. Given his infinity, God necessarily creates as essences truths that contradict logical truths and

---

<sup>12</sup> For an at least apparently opposite view refer to Ragland (2005).

might have instantiated all or some of them in actual world. In this sense, any truth achieved by finite human mind is restricted to what humans can conceive, but is not necessarily absolutely truth. And this would include the truths alleged proved in Descartes's *Meditations* such as my existence, God's existence, his veracity, and so on. There would be no guarantee of absolute knowledge by human finite minds. Reason would not be the ensured source of knowledge, but rather the indubitable source of knowledge. Because it is not dubitable, reason has the best credentials for being the source of human knowledge. This seems to be what Descartes means when he says:

What is it to us that someone may make out that the perception whose truth we are so firmly convinced of may appear false to God or an angel, so that it is absolutely speaking, false? *Why should this alleged 'absolute falsity' bother us, since we neither believe in it nor have even the smallest suspicion of it?* [AT VII: 144; CSM II: 103; emphasis added].

Judging by the way in which Descartes preserves divine veracity in relation to the senses, it does not seem implausible that God, though truthful, has created an actual world resistive to the human intellect.

If my reading is correct, that is, if Descartes's free creation doctrine involves that human mind does not reach absolute truth, then either we admit that Descartes's whole metaphysical project fails or we move Descartes away from a project in which he is pursuing the guarantee of reason as the source of absolute truth (that is, truth from the point of view of the creator) towards a project that involves considering reason only as the best candidate for achieving the truth relative to humans limitation. If the latter, God's role in Cartesian system has to be revised. But this discussion goes beyond the scope of this paper.

#### REFERENCES

- Beysade, M. (1990). *Méditations métaphysiques* (trad.), Librairie Générale Française;
- Carriero, J. (2009). *Between Two Worlds: A Reading of Descartes's Meditations*. Princeton and Oxford: Princeton University Press.;
- Croning, T. J. (1960). Eternal Truths in the Thought of Descartes and of His Adversary. *Journal of the History of Ideas*, 4, 553-559;
- Frankfurt, H. (1970). *Demons, Dreamers and Madmen*. New York: Bobbs-Merril;
- Frankfurt, H. (1977). Descartes on the Creation of the Eternal Truths. *The Philosophical Review*, 86. 36-57;
- Gilson, E. (1930). *Études sur le rôle de la pensée médiévale dans la formation du système Cartésien*. Paris: Vrin;
- Gueroult, M. (1953). *Descartes selon l'ordre des raisons*, 2 vols. Paris: Aubier;
- Gouhier, H. (1978). *La pensée métaphysique de Descartes*. Paris: Vrin;
- Greene, M. (1978). *Descartes*. Indianapolis: Hackett;
- Hoffman, P. (2002). Descartes's Theory of Distinction. *Philosophy and Phenomenological Quarterly*, 64, 57-78;
- Ishiguro, H. (1986). The Status of Necessity and Impossibility in Descartes. In A. O. Rorty (Ed.). *Essays on Descartes' Meditations* (pp. 459-472). California: University of California Press;
- Kaufman, D. (2003). Divine Simplicity and the Eternal Truths in Descartes. *British Journal for the History of Philosophy*, 11, 553-579;
- Kaufman, D. (2005). God's Immutability and the Necessity of Descartes's Eternal Truths. *Journal of the History of Philosophy*, 43, 1-19;
- Laporte, J. (1945). *Le Rationalisme de Descartes*. Paris: Presses Universitaires de France;
- Machamer, P., & McGuire, J.E. (2009). *Descartes's Changing Mind*. Princeton and Oxford:



Princeton University Press;

Nolan, L. (1997). The Ontological Status of Cartesian Nature. *Pacific Philosophical Quarterly* 78, 169–194;

Normore, C. (1991). Descartes's Possibilities. In G.J.D. Moyal (Ed.), *René Descartes – Critical Assessments*, Vol. III (pp. 68–83). Routledge;

Osler, M. (2004). *Divine Will and the Mechanical Philosophy*. Cambridge: Cambridge University Press;

Pessin, A. (2010). Divine Simplicity and the Eternal Truths: Descartes and the Scholastics. *Philosophia*, 38, 69–105;

Ragland, C.P. (2005). Descartes on Divine Providence and Human Freedom. *Archiv für Geschichte der Philosophie*, 87(2): 159–188;

Rozemond, M. (2008). Descartes's Ontology of the Eternal Truth. In Gideon Yaffe, David Owen, and Paul Hoffman (Eds.), *Contemporary Perspectives on Early Modern Philosophy: Essays in Honor of Vere Chappell* (pp. 41–63). Peterborough, Ontario: Broadview Press;

Schmaltz, T. M. (1991). Platonism and Descartes' View of Immutable Essences. *Archiv für Geschichte der Philosophie*, 73, 2, 129–170;

Skirry, Justin (2004) Descartes's Conceptual Distinction and its Ontological Import, *Journal of the History of Philosophy*, 42, 121–144;

Walski, G. (2003). The Cartesian God and the Eternal Truths. In D. Garber & S. Nadler (Eds.). *Oxford Studies in Early Modern Philosophy*, vol. 1 (pp. 23–44), Oxford: Clarendon Press.;

Wilson, M. (1978). *Descartes Ego Cogito, Ergo Sum*, ed. T. Honderich, London, New York, Routledge & Kegan Paul.

---

MARCOS ANDRÉ GLEIZER

Universidade do Estado do Rio de Janeiro

mgleizer@globo.com

---

# SPINOZA ON METAPHYSICAL DOUBT AND THE “CARTESIAN CIRCLE”

---

## *abstract*

*This article offers an analysis and defense of the solution proposed by Spinoza to the “Cartesian circle” problem. Taking into consideration Spinoza’s sound analysis of the epistemic conditions of an authentic doubt, it will try to show, against the interpreters who maintain that Spinoza’s most explicit and consistent solution fails, that his solution offers a perfectly coherent account of the self-justification of the objective value of reason. I will also briefly indicate the intimate connection existing between Spinoza’s solution to the “Cartesian circle” problem and his conception of truth as its own standard*

---

## *keywords*

*Spinoza, Descartes, metaphysical doubt, definition of truth, criterion of truth, self-justification*

---

My aim in this paper is to analyze the solution proposed by Spinoza to the problem traditionally known as the “Cartesian circle”. Arnauld, in his *Fourth Objections* to Descartes’ *Metaphysical Meditations*, formulated this problem in the following way:

We are sure that what we clearly and distinctly perceive is true only because God exists. But we can be sure that God exists only because we clearly and distinctly perceive this. Hence, before we can be sure that God exists, we ought to be able to be sure that whatever we perceive clearly and evidently is true. (AT VII 214)

- 1. Significance of the problem** Once we consider that the set of things that we conceive clearly and distinctly contains the set of rational ideas and principles, the philosophical problem involved in the “Cartesian circle” problem takes the form of the following question: How is it possible a self-justification of the objective value of reason? Finding a solution for this problem consists in showing that there is no vicious circle implied in such self-justification, thus clarifying its meaning and possibility. Of course, the task of legitimizing the truth claims of rational knowledge is justifiable only insofar as the skeptic challenge to such claims is taken seriously. In other words, it is necessary to justify reason only because challenging it was previously taken as possible.
- 2. That Spinoza dealt with this problem** Spinoza was well aware of this problem. Not only did he expound it in the Prolegomenon of his book on *Descartes’ Principles of Philosophy*, but he dealt with it again both in his *Treatise on the Emendation of the Intellect* (§79) and – what is less noted and explored by his commentators – in his *Theological-Political Treatise* (chapter VI with its note).<sup>1</sup> The simple presence of this problem in two works where Spinoza elaborates or expounds his own philosophical thought, and not Descartes’, is already a strong indication that he was not just aware of it as a “historian of Cartesian philosophy”, but also that he faced it as a true problem for him as a philosopher. Now, what does a philosopher who claims that “truth is the standard both of itself and of the false” (EIIP43S, GII/124) have to say about this problem? Spinoza is generally seen as the very paradigmatic case of philosophical dogmatism found in modern times; as someone who, by affirming the identity between true ideas and certainty, would have suppressed the problem of

---

<sup>1</sup> These works will be henceforth referred to as *PP*, *TdIE* and *TTP*, respectively. All emphasis in quoted sentences was added.

certainty, instead of really discussing it.<sup>2</sup> His absolute rationalism seems thus to be based only on an act of blind faith in the value of reason.

Spinoza's complex theory of truth certainly entails that true ideas are, through their intrinsic property, directly recognizable without the need of any sign.<sup>3</sup> Nevertheless, it is wrong to infer from that, that truth is immediately recognized or attained without any intellectual effort, or that no obstacles can hinder its self-manifestation. Indeed, Spinoza insists in many passages that the force of imaginative prejudices can blind our thought, hinder our apprehension of the distinction between true and false ideas, and raise doubts.<sup>4</sup> Moreover, the fact that true ideas contain two distinct properties (*adaequatio* and *convenientia*), whose necessary connection can only be proved when we have an adequate knowledge of God, entails that, as long as we do not possess explicitly this knowledge implicitly involved in every idea (EIIIP46, GII/127), we can doubt everything.<sup>5</sup> Thus, even if it is true that struggling with skepticism is not Spinoza's chief philosophical concern, this does not mean that he did not recognize the relevance of this problem in his system.

It is also true that doubt does not and cannot perform any kind of methodological function in Spinoza's epistemological thought. Nonetheless, this does not mean that the state of doubt – understood as an effectively experienced mode of thought, a modality of non-true idea that prevents the full self-manifestation of truth –, does not receive from Spinoza a very precise analysis, one that explains exactly in which conditions it necessarily takes place, and how we can legitimately remove it. While investigating this, Spinoza never avoids or denies the possibility of a genuine doubt about the truth of clear and distinct ideas. He faces this radical possibility and offers a solution. My intention is to investigate this solution.

---

2 Cf. Hubbeling, 1967, p. 35, and Hamelin, 1984, p.102.

3 The full understanding of Spinoza's concepts of truth and certainty requires a thorough interpretation whose presentation goes beyond the scope of this paper. I developed this interpretation in Gleizer (2017), where I attempted to show that Spinoza's originality consists in suppressing the false opposition between the conceptions of truth as coherence and as correspondence. According to his theory, *adaequatio* (coherence) and *convenientia* (correspondence) are two complementary aspects necessarily involved in the concept of truth, so that, for an idea to be true, it must fulfill two conditions: be adequate and agree with its object. Adequacy, being an intrinsic property that takes part in the very definition of true ideas, also functions as what manifests their truth; not, however, as a sign, for signs are variable, arbitrary, and have no intrinsic connection with what they signify (see *TTP* chapter 2, and EIIIP18S). Nevertheless, since the necessary connection between the intrinsic and the extrinsic properties constitutive of true ideas can only manifest itself when we grasp some of the consequences that follow from the adequate idea of God, as long as this idea and these consequences are not grasped, doubt can arise.

4 See, among other passages, *TdIE* §45 (GII/17) and §47 (GII/18), *TTP* Preface (GIII/8) and chapter XV (GIII/180). Even in the *Ethics*, we find passages where Spinoza insists on this point. The most striking one is found in the scholium to the very proposition in which Spinoza establishes that truth is its own standard (EIIIP43S): “for to have a true idea means nothing other than knowing a thing perfectly, or in the best way. And of course no one can doubt this *unless he thinks that* an idea is something mute, like a picture on a tablet, and not a mode of thinking, viz. the very [act of] understanding” (GII/24). The expression “unless he thinks that” introduces precisely the condition under which a doubt concerning what does it mean to have a true idea becomes possible, this condition being the presence of an imaginative belief concerning the nature of ideas.

5 See *TTP* Chapter IV: “Now since all our knowledge, and the certainty that banishes every possible doubt, depend solely on the knowledge of God – because firstly, without God nothing can be or be conceived, and secondly, everything can be called into doubt as long as we have no clear and distinct idea of God – it follows that our supreme good and perfection depends solely on the knowledge of God.” (GIII/59-60)

**3. On the conditions of possibility of doubt according to Spinoza**

To do so, it is first required that we expose Spinoza’s analysis of the conditions of possibility of doubt, as it is undertaken both in the *Ethics* and in the *TdIE*. Inverting the chronological order between these two works, let us begin by briefly recalling the main elements of the analysis proposed in the *Ethics*. This analysis can be split in two distinct moments: one negative and the other positive.

**3.1. Negative moment**

The negative moment consists in a detailed refutation of the foundations of the Cartesian theory of judgment and, hence, of its account of doubt and error. I shall not attempt to reconstruct Spinoza’s argument here, but his critique might be synthesized in the following theses:

- 1.1. “In the mind there is no absolute, or free, will” (EIIIP48, GII/129).
- 1.2. “in the mind there is no absolute faculties” (EIIIP48S, GII/129).
- 1.3. faculties of the mind are just entities of reason (EIIIP48S: “...these and similar faculties are either complete fictions or nothing but metaphysical beings, or universals...”).
- 1.4. “in the mind there is no volition, or affirmation and negation, except that which the idea involves insofar as it is an idea” (EIIIP49, GII/130). According to this important thesis, there is no difference between the act of considering a thought-content and the act of giving assent to it, between the cognitive process by which a propositional content is considered and that by which this content is subsumed to a volitional act. This amounts to saying that perceiving is the same as affirming something to be the case, or that propositions spontaneously take place as beliefs in the mind.

Thus,

- 1.5. the suspension of judgment is not an *act* of free will, through which the mind would be able to suspend assent to what is perceived by the understanding (EIIIP49S, GII/134).

**3.2. Positive moment: Spinoza’s explanation of the suspension of judgment**

The exclusion of free will renders the analysis of the epistemic conditions of doubt very precise. The human mind, understood as a ‘spiritual automaton’, is submitted to the necessary laws regulating the logic of mental life, laws that are not disturbed by the presence of any “absolute power of decision” that would take us out of the realm of natural causality and explanation.

Thus, the positive moment shows that, rather than being an *act*, the suspension of judgment is a passive *state* in which the mind *necessarily* finds itself whenever occurs a certain situation of conflict between *two ideas*. This state consists in a vacillation between these ideas.<sup>6</sup> Since every idea involves an affirmation, doubt does not consist in the suppression of the idea’s affirmation, but in a *logical instability* brought about by the presence of a different and conflicting affirmation. This *logical instability* means that the mind, when in the epistemic situation of doubt, is not capable of arriving at a definite conclusion about the object of doubt. What are the characteristics of ideas that conflict in doubt? According to EIIIP49S, doubt takes place when the mind “sees that it does not perceive something adequately”, that is, when it

---

<sup>6</sup> In EIIIP17S, Spinoza asserts that between doubt and affective vacillation there is merely a difference of degree, and not of nature: “This constitution of the mind which arises from two contrary affects is called vacillation of mind (*fluctuatio Animi*), which is therefore related to the affect as doubt is to the imagination (see EIIIP44S); nor do vacillation of mind and doubt differ from one another except in degree.” (GII/153)

perceives that it has an inadequate idea. The doubtful idea is thus specified as inadequate, and the conflicting idea that raises doubt is characterized as the perception of a perception, that is, as a reflective judgment.<sup>7</sup>

Let us turn now to the more detailed version of the analysis presented in the *TdIE*.

Here, Spinoza introduces his analysis of the *idea dubia* by way of the distinction between the merely verbal doubt and the authentic one (*vera dubitatione in mente*, §77). This distinction is a particular case of a broader distinction between what we can say and what we can think.

Not only must we not confuse words and ideas, verbal affirmations and mental affirmations, but neither must we believe the former to be always expressions of the latter.<sup>8</sup> Merely verbal doubt is a linguistic utterance that, because it does not satisfy the conditions of an authentic doubt, does not express any thought.

According to §78, two conditions must be fulfilled if doubt is really to take place:

1. It is necessary, as in the *Ethics*, that two ideas be given: the idea of the object of doubt (idea *p*) and the idea that makes us be in doubt, that is, the reason, cause or motive of doubt (idea *q*). This means that no idea is doubtful in itself, but is rendered doubtful by its relation to another idea.<sup>9</sup>
2. It is necessary that the reason for doubting (idea *q*) be not clear and distinct, but confused. In Spinoza's words:

There is no doubt in the soul, therefore, through the thing itself concerning which one doubts. That is, if there should be only one idea in the soul, then, whether it is true or false, there will be neither doubt nor certainty (...). But doubt will arise through another idea which is not so clear and distinct that we can infer from it something certain about the thing concerning which there is doubt. That is, the idea that puts us in doubt is not clear and distinct. (*TdIE* §78, GII/29-30)

According to this passage, and unlike the one quoted from the *Ethics*, there is no indication as for the character confused or not of the idea of the object of doubt (idea *q*). The possibility seems thus open for its being both confused or clear and distinct. Neither are there indications of whether the idea that throws us in doubt (idea *q*) originates from a reflective level or not. On the other hand, it is stressed that the reason for doubting is *necessarily confused*, for if it were clear and distinct, we could infer from it something certain about the object of doubt, in which case there would be no doubt. In Spinoza's words:

Doubt is nothing but the suspension of the mind concerning some affirmation or negation, which it would affirm or deny if something did not occur to it, the ignorance of which must render its knowledge of the thing imperfect. (*TdIE* §80, GII/30)

---

<sup>7</sup> However, the conflicting ideas causing doubt do not seem to have necessarily a reflective origin, for in EIIP44S, where Spinoza explains the origin of our inadequate idea of things as contingent, he offers an example of vacillation caused by a conflict of imaginative ideas coming from a pre-reflexive level.

<sup>8</sup> On this respect, see EIIP49S, specially the following passage: "those who confuse words with the idea, or with the very affirmation the idea involves, think that they can will something contrary to what they are aware of, when they only affirm or deny with words something contrary to what they are aware of." (GII/132)

<sup>9</sup> Descartes already noted this relational condition of doubt in the (K) item of his *Responses to the Seventh Objections*: "It should be noted that throughout he [Bourdin] treats doubt and certainty not as relations of our thought to objects, but as properties of the objects which inhere in them for all time. This means that if we have once realized that something is doubtful, it can never be rendered certain." (AT VII 473). The word 'object' here refers to mental items, such as ideas, perceptions or judgments.

The reason for doubting, therefore, is a confused, imperfect and inadequate idea, and not a certain one. In other words, doubt (and hence the skeptical position) does not presuppose certainty: its roots lie in *ignorance*.<sup>10</sup> Doney (1975) thus synthesizes the elements of this analysis:

If someone claims that *p* is doubtful because it is possible that *q*, he implies that he does not know that *q* and also that he does not know that not-*q*. (p. 145).

This formulation indicates that the mere fact of considering something as possible lies in our ignorance about the cause (or reason) that necessarily posits or excludes it, and stresses also that to doubt *p*, another idea, different from *p*, is needed, namely, the idea *q*. To claim that *p* is doubtful simply because it is possible that not-*p* is not the same as offering a reason for doubt: it is merely a *gratuitous and groundless affirmation* of the possibility of not-*p*. Therefore, according to Spinoza, for every proposition *p*, if the only reason for doubting *p* is that “maybe not-*p*” (it is possible that not-*p*), then we have no reason for doubting it. Affirmation of doubt, in such case, becomes something merely verbal.

It is important to lay emphasis on the inadequate character of the reason for doubting and on its representation of things as possible. Affirming that reasons for doubt are necessarily inadequate or confused amounts to affirming that they are originated in the imagination and never in reason. This means that it is not reason that casts doubt on itself, that doubt does not spring from a disagreement between reason and itself. According to Spinoza, it is characteristic of rational knowledge to conceive things as necessary, while it is characteristic of imagination to consider them according to the modal categories of possibility and contingency. Spinoza understands these two modalities epistemically, in other words, they only express our ignorance of what renders things necessary or impossible.<sup>11</sup> According to the analysis of fictitious ideas proposed in the *TdIE*, imagination takes advantage of this absence of intellection to engender, through passive associations of confused ideas, its fictions, hypotheses and suppositions about the essence and existence of things. We can form fictions only insofar as we consider some object as possible, that is, while its necessity or its impossibility is unknown to us. Thus, the power of forming fictions is inversely proportional to that of understanding (*TdIE* §58, GII/22), so that, after we have intellectually grasped the necessity or impossibility of an object, we can no longer mentally ascribe some predicates to it.<sup>12</sup> Nevertheless, as long as this understanding has not yet happened, “we think that the things we more easily imagine are clearer to us, and think we understand what we imagine” (*TdIE* §90, GII/33). In other words, imagination, left to itself, takes itself spontaneously for a true knowledge: it is not “*index sui*”. Reasons for doubting are necessarily confused, but their

---

<sup>10</sup> Descartes also pointed out this aspect, emphasized by Spinoza as a *necessary* condition of doubt, in the (M) item of his *Responses to the Seventh Objections*. Against Bourdin’s objection that good and strong reasons for doubting should be certain, Descartes writes: “There may be reasons which are strong enough to compel us to doubt, even though these reasons are themselves doubtful, and hence are not to be retained later on, as I have just pointed out. The reasons are strong so long as we have no others which produce certainty by removing the doubt” (AT- VII- 473-474). Both for Spinoza and Descartes, a given reason for doubting has force only insofar as contrary intellectual evidence excluding it is not present, and not by virtue of any intellectual evidence it itself might have.

<sup>11</sup> Cf. *TdIE* §53, EIP33S1.

<sup>12</sup> Cf. *TdIE* §53 (GII/19-20): “Here I ask, what does such an idea concern? I see that it concerns only possible, and not necessary or impossible things. ... If its necessity or impossibility, which depends on external causes, were known to us, we would have been able to feign nothing concerning it”; and *TdIE* §58 (GII/22): “after we know the nature of body, we cannot feign an infinite fly, or after we know the nature of the soul, we cannot feign that it is square, though there is nothing that cannot be put in words”.

confusion is not necessarily manifest, and this accounts for their initial plausibility. It is thus clear that an imaginative idea will work efficiently as a reason for doubting only insofar as the necessity or impossibility of what it displays “as possible” escapes us: that is, as long as there is no intellectual evidence (an apodictic apprehension) positing or excluding necessarily its object. It is only during this period of ignorance that the content of fictions, hypotheses and other kinds of imaginative ideas that drive the mind to a state of doubt can present themselves “as possible”.

Since every reason for doubting is thus an inadequate, confused idea that exhibits its object under the modality of possibility, how can we remove it? In order to achieve this, it is *necessary and sufficient* that we form an adequate idea that replaces the inadequate idea that raises doubt. If we are able to do this, we will know, to go back to Doney’s formulation, whether *q* is necessary or impossible, and we will either necessarily affirm or necessarily deny *q*. Thus, doubt about *p* will be excluded, for we will be able to infer from *q* something certain about *p*. With all that in mind, let us now turn to Spinoza’s exposition of the problem of the “Cartesian circle”.

In the prolegomenon to *PP*, Spinoza thus presents the objection of circularity raised against Descartes:

But before we finish, it seems we must satisfy those who make the following objection. Since God’s existence does not become known to us through itself, we seem unable ever to be certain of anything (...). For we have said that everything is uncertain so long as we are ignorant of our origin, and from uncertain premises, nothing certain can be inferred. (*PP* p. 236, G1/146)

This objection is based on the following theses:

1. “God’s existence does not become known to us through itself”, that is, it must be an object of demonstration.
2. “Everything is uncertain so long as we are ignorant of our origin” (that is, of God’s existence and veracity).
3. “From uncertain premises, nothing certain can be inferred”.

Once we accept these three theses, we must infer that “we seem unable ever to be certain of anything”.

Having presented the objection, Spinoza goes on to expose his interpretation of the Cartesian response such as it is presented in the *Principles of Philosophy*, Part I, Article 13, in the *Responses to the Second Objections*, point 3, and in the end of the *Fifth Meditation*. According to Spinoza, Descartes’ response consists in weakening the second thesis, inasmuch as, in those texts, Descartes limits the scope of doubt to remembered evidences. Present evidence is beyond doubt and requires no divine guarantee, which has to do exclusively with the science of the conclusions that can be separated from their premises (and only when they are actually so). Propositions that require no proof (first principles, axioms or common notions), and demonstrations while they are being accomplished (present evidence), being unaffected by metaphysical doubt, do not require God’s guarantee. As Spinoza has it:

From the fact that we do not yet know whether the author of our origin has perhaps created us so that we are deceived even in those things that appear most evident to us, we cannot in any way doubt the things that we understand clearly and distinctly either through themselves or through reasoning (so long, at any rate, as we attend to that

#### **4. Spinoza’s presentation of the problem of the “Cartesian circle”**

##### **4.1. Presentation of the objection**

##### **4.2. Interpretation of the Cartesian Response**



reasoning). We can doubt only those things that we have previously demonstrated to be true, and whose memory can recur when we no longer attend to the reasons from which we deduced them and, indeed, have forgotten the reasons. (PP p. 236, GII/147)

Now, since Descartes admits the possibility of proving the existence of God, as long as we remain attentive to all the premises (particular present evidences) from which this conclusion derives, it is possible to know this existence with certainty.

It is not my intention to discuss here the accuracy of Spinoza’s interpretation of the Cartesian solution. My aim is to analyze Spinoza’s other response – the specifically spinozistic one –, which is found also in the *TdIE* §79 and the *TTP*, and in which Spinoza speaks on his own behalf. This response is presented as an alternative to the first, since this one “does not satisfy some people” (GII/147). It is noteworthy that Spinoza never states explicitly that the Cartesian response does not satisfy himself, although in other places of his work he clearly expresses his dissatisfaction with some of Descartes’ demonstrations. Does he see this solution as a valid one? We shall see that this is not so.

#### 4.3. Rejection of the Cartesian solution

Indeed, the simple fact that Spinoza presented his response as an *alternative* to the first, and not as a reconstruction of it, is enough to indicate that he shared that dissatisfaction. On the other hand, when we consider the introductory passages where Spinoza refers to the metaphysical doubt, we verify that they allude to passages of Descartes’ works where the scope of doubt seems not to be limited to remembered evidences, but to include also present evidences, as well as truths that are simpler than mathematical truths.<sup>13</sup>

However, it is not only the references employed by Spinoza that seem to suggest his dissatisfaction with a solution that seeks to limit the scope of doubt. This is confirmed by the way the problem is presented in the *TdIE*, in the *TTP* and in the very sequence of the *PP*. In the *TdIE*, Spinoza expresses in the following way the possibility of the metaphysical doubt:

From this it follows that, only so long as we have no clear and distinct idea of God, can we call true ideas in doubt by supposing that perhaps some deceiving God exists, who misleads us even in the things most certain (*maxime certis*). (*TdIE* §79, GII/30)

No restrictions to the scope of doubt are formulated here: there is no mention of any

---

<sup>13</sup> The passages are the following:

a) Doubt concerning all things: “nevertheless he discovered a reason for doubting [mathematical truths]: for (...) deeply rooted in his mind was an old opinion, according to which there is a God who can do all things and by whom he was created such as he was. Perhaps this God had made him so that he would be deceived even about those things that seemed clearest to him” (PP p. 232, GI/143). This passage seems to refer to the end of the *First Meditation* where Descartes states that: “And yet firmly rooted in my mind is the long-standing opinion that there is an omnipotent God who made me the kind of creature that I am. (...) What is more, since I sometimes believe that others go astray in cases where they think they have the most perfect knowledge, may I not similarly go wrong every time I add two and three or count the sides of a square, or in some even simpler matter, if that is imaginable?” (AT- VII- 21). Even the simplest propositions and present evidence are here said to be within the scope of metaphysical doubt.

b) Liberation from all doubts: “For when he discovers that there is a most perfect being, (...) with whose nature being a deceiver is incompatible, then that reason for doubting which he had because he was ignorant of his cause will be removed. (...) Hence neither Mathematical truths nor any of those that seem most evident to him can be at all suspected” (PP p.235, GI/145). This passage seems to refer to the *Third Meditation*, where Descartes, after having placed the *cogito* along with mathematical propositions and axioms, concludes by saying that: “But, in order to remove even this slight reason for doubt [the metaphysical reason for doubt], as soon as the opportunity arises I must examine whether there is a God, and, if there is, whether he can be a deceiver. For if I do not know this, it seems that I can never be quite certain about anything else” (AT- VII- 36).

exemption on the part of simple propositions, common notions or present evidence. What is being stated is simply that, as long as we have no knowledge of our origin, and consider the hypothesis of a deceiving God, we can doubt “the things most certain”.

Nor does Spinoza make any restriction to the scope of doubt in the *TTP*. He says, in Chapter IV:

...all our knowledge, and the certainty which removes every doubt, depend solely on the knowledge of God: firstly, because without God nothing can exist or be conceived; secondly, because so long as we have no clear and distinct idea of God we may remain in universal doubt. (*TTP*, Ch. IV, GIII/59-60)

Finally, as the prolegomenon to *PP* goes on to show, Spinoza opposes himself to the Cartesian solution by clearly stating that the attention given to the demonstration of a proposition does not make us absolutely certain of that proposition:

We can not be absolutely certain of anything, except our own existence, *even though we attend properly to its demonstration*, so long as we have no clear and distinct concept of God that makes us affirm that he is supremely veracious (*PP*, p. 237, GI/148)

The passages quoted above show clearly that, for Spinoza, mere attention given to present evidence does not accord to this evidence an absolute certainty. What Spinoza seeks to call in question through the notion of “attention” is the legitimacy of assimilating *persuasion* (irresistible inclination to assent caused by the presence of evidence) to *certainty*. This is not to say that he does not acknowledge the *factual impossibility* of doubting present evidence. On the contrary, Spinoza acknowledges it inasmuch as, according to him, every idea (be it clear and distinct or mutilated and confused) involves the affirmation of its own content, that is, inasmuch as he makes no distinction between perceiving and affirming. In order to doubt, it is necessary to turn away from the idea of the object of doubt and consider the reason for doubting. Evidence can only be shaken retrospectively. As Spinoza puts it in the prolegomenon to *PP*: when we turn our minds’ eyes towards the nature of the triangle, we are “compelled to infer” that its three angles are equal to two right angles; but when our mind faces the hypothesis of the deceiving God, we can cast doubt on that property. Thus, to doubt it is necessary to turn the mind’s eyes toward the reason for doubting. Nevertheless, the *impossibility of fact*, or *psychological incapacity*, of doubting present evidence while considering it attentively is not to be confused with the *logical exclusion* of the metaphysical reason for doubting, and does not amount to establishing the *right* to know things as they are in themselves. The irresistible inclination to believe that *p* is not the same as the logical impossibility of doubting *p*. Besides, the *moment of doubt* must not engender a confusion concerning the *object of doubt*. If doubt is always retrospective, this does not mean that the object of doubt is not the value of present evidence as the criterion for truth. Therefore, as long as the movement of thought by means of which we psychologically and momentarily escape from the action of metaphysical doubt does not coincide with the movement through which the latter is logically destroyed, present evidence cannot be qualified as absolutely certain.

#### 4.4. Reason for this rejection

Spinoza, however, makes a noteworthy exception for the *cogito*. He not only qualifies it as *absolutely certain*, but takes it to be the very *paradigm of certainty*. In the beginning of the exposition of his response to the objection, Spinoza reminds the reader that:

When we previously discussed the certainty and evidence of our existence, we saw that we inferred it from the fact that, wherever we turned our attention (...), we came upon

#### 4.5. Exception of the Cogito

no reason for doubting that did not by itself convince us of our existence. (PP, p. 236, GI/147)

Moreover, he affirms that after we have formed a clear and distinct idea of God, mathematical truths will be in the same situation as the *cogito*:

Wherever we direct our attention in order to doubt some one of them, we shall come upon nothing from which we must not instead infer that it is most certain – as happened concerning our existence. (PP, p. 237, GI/148)

The *cogito* thus emerges as the paradigm of an epistemic situation in which the reasons for doubting are logically neutralized, being transformed into, or substituted by, reasons for believing. This clearly indicates the strategy of justification that Spinoza will adopt, which consists in *suppressing the obstacle that hinders us from believing in the truth of our clear and distinct ideas*. It must be stressed, however, that, if the *cogito* is taken, in that text, as the paradigm of certainty, Spinoza does not understand it as furnishing the inescapable point of departure or the sole premise from which it is possible to develop an unshakeable proof of God’s existence. He clearly affirms that, in order to suppress metaphysical doubt, we just need to arrive at the clear and distinct idea of God, “however we have acquired it” (PP, p. 238, GI/148).<sup>14</sup> According to this analysis, it seems clear that Spinoza is not satisfied with the solution that tries to attack thesis (2) in order to limit its scope. As long as the fiction of a deceiving God is considered “as possible”, he accepts the maximum expression of metaphysical doubt.

#### 4.6. Analysis of the three theses

In order to move forward into the properly spinozistic response, it is necessary to examine how Spinoza sets the problem, that is, how exactly he understands the three theses enunciated above.

##### 4.6.1. Thesis 1

The first thesis (“God’s existence does not become known to us through itself”) does not seem to pose any interpretive difficulty. We find, for example, the following passage in the *TTP*, Chapter VI:

*As God’s existence is not known through itself*, it must necessarily be inferred from notions so firmly and incontrovertibly true, that no power can be postulated or conceived sufficient to impugn them (*TTP*, Ch. VI, GIII/84)

---

<sup>14</sup> Curiously, no exception for the *cogito* is found either in the *TdIE* §79 or in the *TTP*, and even by the end of the Prolegomenon to the *PP*, where Spinoza presents the objection of the “circle” in a syllogistic form, he concedes the major premise of this syllogism formulated in a totally general way: “we can be certain of nothing before we have a clear and distinct idea of God” (GI/149). However, Spinoza makes some implicit references to the certainty of the *cogito* in the *TdIE* (§54, GII/20, and §58, GII/22), and even uses it as a kind of last resort to refute those skeptics who, being afraid to confess that they exist if they say that they know nothing, are like “automata, completely lacking a mind” (§§47-48, GII/18). The problem, therefore, is to determine the specificity of its certainty. Although it resists the metaphysical doubt, it does not destroy the deceiving God hypothesis. As far as the *cogito* is concerned, reiteration of doubt strengthens its certainty. In this case, the metaphysical reason for doubting is neutralized by logico-pragmatic reasons, but the internal inconsistency it implicitly involves is not revealed. This is the reason why we must not say that the metaphysical doubt is *destroyed*, but only that its scope of action is limited because it comes across a specific case that *resists* it and repeatedly neutralizes its effect. We may, therefore, distinguish three modalities of neutralization of doubt: 1) present mathematical evidences only *escape* psychologically and momentarily from the action of doubt; 2) the *cogito*’s present evidence logico-pragmatically neutralizes this action, repeatedly *resisting* the assault of doubt without, however, *destroying* it; 3) the present evidence of the idea of God logically *destroys* the reason for doubting, liberating evidences (1) and (2) from the necessity of being constantly reactivated.

The *Ethics*, in its turn, reinforces the need of demonstrating God's existence, this demonstration being precisely the object of EIP11. Contrary to what some interpreters maintain, it is thus clear that Spinoza's solution does not consist in denying the first thesis, affirming an immediate, non-inferential access to the existence of God.<sup>15</sup>

As for the second thesis, we have already seen that Spinoza accepts its most radical expression *as for its scope*. There is, however, a second aspect of it that he seems to have altered. Instead of affirming that we can doubt everything *as long as we have no knowledge of God's existence and veracity*, he affirms that we can doubt everything as long as *we have no clear and distinct idea of God*:

We can be certain of nothing – not, indeed, so long as we are ignorant of God's existence (for I have spoken nothing of this) – but as long as we do not have a clear and distinct idea of him. (*PP*, p. 238, GI/149)

Spinoza establishes here a distinction between having knowledge of God's existence and veracity and having a clear and distinct idea of God. What does this distinction mean? What is Spinoza's purpose in making it?

He does not mean to hold that knowing that a veracious God exists is not important for the problem at hand, since he clearly affirms in the *TTP* that: "We doubt God's existence, and consequently we doubt everything, so long as the idea we have of God himself is not clear and distinct, but confused." (Chapter VI, marginal note, GIII/84)

The distinction introduced by Spinoza is intended as a way of indicating the deep root from which doubt springs, namely, the absence of a clear and distinct idea of God's nature and the presence of a confused one. We can doubt clear and distinct ideas only insofar as we can doubt the necessary existence of a veracious God. Nevertheless, we can doubt this necessary existence only insofar as we do not have an adequate idea of God's essence, but only a confused one. It is only by virtue of this confused idea that we are as easily able to affirm that God exists as to affirm that he does not exist, and also to affirm that he is a deceiver as to affirm that he is not. This confused idea, displaying God as having just a *possible* existence and as being a *possible* deceiver, is thus what renders metaphysical doubt effective. As Spinoza affirms in the prolegomenon to *PP*:

**4.6.2. Thesis 2:  
the distinction  
between knowing  
the existence of  
God and having a  
clear and distinct  
idea of him**

---

<sup>15</sup> The interpretation according to which the validity of Spinoza's solution to the "Cartesian circle" problem depends essentially on an immediate recognition of God's perfection was supported by Martha Bolton (1985, see, particularly, pp. 382-384). It is true that, in the EIP8S2, Spinoza affirms that: "if men would attend to the nature of substance, they would have no doubt at all of the truth of P7. Indeed, this proposition would be an axiom for everyone" (GI/50), and, consequently, would be a *per se nota* truth. Nevertheless, the use of the conditional and the whole passage reinforces an important point made in the *TdIE* §§43-46, according to which the presence of prejudices prevents us from following immediately the proper deductive order, hinders what is evident by itself to be immediately evident for us, and forces our mind to prepare a way to apprehend what is *per se nota*. The problem of doubt arises only insofar as we have imaginative prejudices, but we naturally have them. In this aspect, Spinoza agrees with Descartes (although he completely disagrees as for the correct method to deal with prejudices). There is, however, an important difference between them concerning knowledge of God's existence by minds already freed of prejudices. While Descartes maintains that this existence would be known *without proof* by these minds, through the mere contemplation of God's nature (see, specially, the fifth postulate (AT- IX- 126-127) and the brief commentary to the a priori proof (AT- IX- 129), both presented in the Geometrical Exposition that follows the *Second Replies*) – what amounts to deny, for them, the first thesis –, for Spinoza, when it is a question of knowing God's existence from its essence, even these minds have to prove the reality of God's definition as an absolutely infinite substance (EIdéf.6), that is, since the "eyes of the mind, by which it sees and observes things, are the demonstrations themselves" (EVP23S), they need to demonstrate the consistency of this definition through the genetic construction of this concept. This is precisely what is done in the first ten propositions of the *Ethics*, thus arriving, according to the *TdIE*'s injunction, "as soon as possible" (§99, GII/36) to God's essence.

For to someone who does not have a true idea of God (...), it is just as easy to think that his author is a deceiver as to think that he is not a deceiver. Similarly, for one who has no idea of a triangle, it is just as easy to think that its three angles are equal to two right angles, as to think that they are not. (PP, p. 237, GI/147)

As we saw above, for doubt to take place it is necessary that two ideas be present, and also that the idea that causes us to be in doubt be inadequate, that is, it must involve the ignorance of something that makes it imperfect, exhibiting its object under the modal categories of possibility and contingency. The deep root of metaphysical doubt is the presence of an inadequate (confused) idea of our origin.<sup>16</sup>

It must also be stressed that Spinoza’s reformulation of the second thesis does not mean that, in order to solve the problem of the circle, it is enough to replace the *demonstration of God’s existence* by a *clarification of His idea*, for what he takes as a clarification of God’s idea is indeed a process brought about by way of demonstrations. As stated in the continuation of the passage from the *TTP* cited above:

But for us to be able to conceive God’s nature clearly and distinctly it is necessary for us to attend to certain very simple notions which they call common, and connect with them those which pertain to the divine nature; then for the first time it becomes evident to us that God exists necessarily and is everywhere, and at the same time that all the things we conceive involve in themselves the nature of God and are conceived through it, and finally, that all those things are true which we conceive adequately. (Ch. VI, marginal note, GIII/84)

It is manifest by this passage that *the clarification of our idea of God is a demonstrative procedure*. Forming a clear and distinct idea of God is nothing but demonstrating, with the help of common notions, what belongs to His nature (existence, omnipresence and veracity). However, if that is so, how can we be certain of the truth of this demonstration? How can we be certain that our demonstration has resulted in true knowledge of God, since we have started from uncertain premises? Indeed, we cannot presuppose the truth of the clear and distinct premises used in the demonstration without presupposing exactly what is at stake. This leads us to the third thesis (“from uncertain premises nothing certain can be inferred”) and to Spinoza’s solution to the “Cartesian circle” problem.

**4.6.3. Thesis 3** The third thesis contains the core of the matter. Does Spinoza accept it without any exception? In the Chapter VI of the *TTP*, he seems to deny it for the specific case of the demonstration of what pertains to God’s nature:

Since God’s existence is not known through itself, it must necessarily be inferred from notions whose truth is so firm and steady that no power can be or be conceived by which they could be changed. *At least they must so appear to us at the time when we*

---

<sup>16</sup> Two aspects present in metaphysical doubt should be carefully distinguished. That which furnishes the *occasion* for the metaphysical doubt to arise is the application of a rational principle, namely, the principle of causality. Only when we ask ourselves for the cause or origin of our mind that doubt *can* arise. However, our ignorance of our origin is not enough for doubt to arise *effectively*: the mere absence of a clear and distinct idea of our origin is not enough to engender doubt. It is necessary to *suppose* that our origin *could* be such as to give rise to a gap between the demands of our rationality and the structure of reality. It is this *supposition* that receives its most radical figuration with the hypothesis of a deceiving God, and thus that provides the confused idea that *effectively* engenders doubt.

*infer God's existence from them*, if we want to infer it from them beyond any chance of doubt. For if we could conceive that the notions themselves could be changed by some power, whatever in the end it was, we would doubt their truth, and consequently also doubt our conclusion, viz., God's existence, nor would we ever be able to be certain of anything. (TTP, Ch. VI, GIII/84)

What exactly is going on here? *Just before the moment we inferred*, from common notions, that God exists – more precisely, that a veracious God exists – it seemed we could conceive some sort of power capable of changing those notions, that is, capable of rendering them false; because then, in that epistemic situation, *we believed it was possible to conceive* a deceiving God whose function was precisely to make us suspicious about the truth of our clear and distinct ideas. Indeed, our confused idea of God exhibited this as a possibility.

However, as Spinoza has it, *at the moment we infer*, from those uncertain notions, that a veracious God exists, it becomes *impossible to conceive* the existence of a deceiving God. According to the third thesis, shouldn't this conclusion be uncertain? Spinoza's answer is no. In this case, not only the conclusion is not uncertain, but it also suppresses the reason for doubting its own premises. The very conclusion hinders us from rendering its premises uncertain.

Is Spinoza consistent in this passage? I would say he is. To blindly accept the third thesis as having an absolutely general validity is to ignore the relational nature of doubt. No idea is uncertain in itself, but is rendered uncertain by the presence of a reason for doubting it. This reason is necessarily a confused idea that is effective as a reason for doubting only insofar as no contrary intellectual evidence is given which reveals the necessity or impossibility of the content it displays “as possible”. Therefore, what stands for a valid reason for doubting in a given epistemic situation may no longer be so in a different one. We also saw that, in order to suppress doubt, it is necessary and sufficient that we form an adequate idea that suppresses the inadequate one responsible for our being in doubt. The *a priori* proof of the necessary existence of a veracious God accomplishes precisely this suppression of the reason for doubting, replacing it by a reason for believing.

In effect, the only reason for doubting the truth of clear and distinct ideas is a confused idea of a deceiving God. Once we form the adequate idea of God's nature, it makes us understand *how* and *why* it is necessary to attribute existence and veracity to God, at the same time revealing to us the *contradiction contained* in the hypothesis of a deceiving God. This fiction, therefore, becomes inconceivable – or, more precisely, that which was already inconceivable, but which we wrongly believed to conceive, is finally showed to be inconceivable. The reason for doubting is thus rejected as a *Chimera*, that is, as a verbal being *lacking all meaning*.<sup>17</sup> Unlike the other rational demonstrations, the demonstrative movement culminating in the conclusion that a veracious God necessarily exists coincides with the movement that reveals the absurd and contradictory character involved in the skeptical hypothesis, logically excluding the only obstacle that hindered us from believing the truth of clear and distinct ideas. To demonstrate that a veracious God creates us is to demonstrate that the hypothesis of a deceiving God is contradictory, that skepticism about reason is not rational.

Thus, once we have formed a clear and distinct idea of God, this idea affects us in such a way as to make it impossible to think that God might deceive us. Therefore, just as with the

---

<sup>17</sup> In his *Metaphysical Thoughts*, Spinoza defines a Chimera as “that whose nature involves an *explicit* contradiction” (Part I, chapter I, note a, G1/233), and says that it is only “a verbal being because it is neither in the intellect nor in the imagination” (Part I, chapter III, G1/241).

paradigmatic case of the *cogito*, wherever we turn the mind’s eyes so as to doubt the common notions from which we began, we can find nothing more that could enable us to infer their uncertainty. This means that whoever forms a clear and distinct idea of God’s nature destroys, in so doing, the only reason he had to doubt the truth of clear and distinct ideas and, *a fortiori*, to doubt the truth of the clear and distinct idea of God.

**4.7. Objection  
1: the clear and  
distinct idea of  
God may have  
been given to us  
by the deceiving  
God**

It could be objected that the *necessity* in which we find ourselves, once we have formed the adequate idea of God, of conceiving him as existing and veracious, and hence the *impossibility* of conceiving the existence of a deceiving God, are nothing but *a necessity and an impossibility for us*, which might not be *a necessity and an impossibility in themselves*. This objection amounts to affirming that it is possible that the adequate idea of the veracious God (a necessary proposition that prevents us from conceiving the existence of a deceiving God) has been given to us by the deceiving God himself, being, therefore, an absolutely false idea. Indeed, no greater subtlety and malignancy could be ascribed to a deceiving God.

It is not clear at all that this objection is indeed conceivable, since it implies that we can still conceive precisely what has been shown to be inconceivable, and that we can still ascribe a meaning to the chimerical expression “deceiving God”. This objection was supported by Michael Della Rocca. He claims that:

The skeptic does not have to concede that, once we have the clear and distinct idea that God is veracious, no reason for doubt of one’s clear and distinct ideas remains. The skeptic can still raise the powerful challenge: why should the world correspond to the deliverances of our most rational thorough investigation? The skeptic would still say that, for all we know, the world fails to match even the system of clear and distinct ideas that includes the distinct and clear idea that there is no deceiving God. The Spinoza of TIE §79 cannot consistently deny this skeptical claim for this claim was invited by Spinoza’s own allowance in that passage that, *at the outset*, we do not know that there is no deceiver. Surely, a truly supreme deceiver could bring about the falsity of even the thorough system of clear and distinct ideas that Spinoza speaks of in TIE §79. This is a *possibility* that Spinoza is in no position to rule out.<sup>18</sup>

His point here seems to be that, once Spinoza provisionally endorses the skeptical reason for doubt, this endorsement will preclude him from ever legitimately removing this doubt. This amounts to accepting Bourdin’s position according to which what is a valid reason to doubt in a certain epistemic context should always remain valid, as if the process of conceptual clarification could not affect it.<sup>19</sup> However, what the production of the clear and distinct idea of God precisely does is to transform the initial epistemic situation, showing that what *seemed possible and thinkable at the outset* (due to our ignorance), involves in fact a hidden contradiction. Once that contradiction is rendered manifest through this process, *it is not*

---

<sup>18</sup> Cf. Della Rocca (, 1994, pp. 33-34). Given the structural similarity between Spinoza’s solution and Gewirth’s interpretation of Descartes that Della Rocca criticizes in his article, the same point is made against the latter (pp. 23-24): “The skeptic could respond to Gewirth’s interpretation of Descartes saying that: ‘even if we do have a clear and distinct idea that God exists and is not a deceiver, *it still seems possible* that this idea and the other clear and distinct ideas apparently validated by it are false – for why should the world correspond to the deliverances of our most thorough rational investigations? [...] Descartes is not in a position legitimately to deny the skeptic’s claim. This is because the skeptic claim is invited by Descartes himself, in particular by his allowance at the beginning of the *Third Meditation* that there may be an omnipotent deceiver”.

<sup>19</sup> See notes 9 and 10.

*possible anymore* to entertain the metaphysical doubt, and hence there is no more reason to doubt the absolute truth of our clear and distinct ideas. What needs to be stressed here is that the plausibility of a reason to doubt depends on its displaying its object as *prima facie* possible, but of course, not everything that seems to be possible is indeed possible.<sup>20</sup>

Furthermore, even supposing that objection to be acceptable, it would not constitute a reason for doubting. As we saw above, it would consist simply in affirming that *p* (“a veracious God necessarily exists”) is doubtful because it is possible that not-*p* (“a deceiving God exists”), and this is nothing but a *gratuitous affirmation* (and not a reason for doubting). Just to affirm the possibility of not-*p* is not the same as exhibiting *another idea* (*idea q*) from which the possibility of this negation would follow. The rules of an authentic doubt would be thus disrespected.<sup>21</sup> In short, the situation would be as follows:

- 1) For every proposition *p*, if the only reason for doubting *p* is that “perhaps not-*p*” (or “it is possible that not-*p*”), then there is no reason for doubting *p*.
- 2) For every proposition *p*, if we understand clearly and distinctly that *p*, then the only reason for doubting *p* would be that perhaps God is deceiving us.

Now, since the proposition that “a veracious God necessarily exists” was clearly and distinctly demonstrated, we must go no further than to substitute *p* for this proposition in order to verify it to be beyond any doubt.

---

20 Della Rocca’s position is reminiscent of Frankfurt’s interpretation of Descartes, although Frankfurt affirms that Descartes’ argument leads to “a conclusion which *excludes the possibility* that there is a demon (or that human existence is a product of chance, or whatever)” (1978, p. 36), and that Descartes establishes that “there is no reasonable ground for doubting that reason is consistent” (1978, p. 39). Frankfurt raises the objection (without calling it a “reason to doubt”) that Descartes’ argument does nothing to show that the consistent system of clear and distinct ideas corresponds to reality. Since he accepts that Descartes gives an explicit account of truth as correspondence (p. 37), how can he avoid transforming this objection into an expression of the metaphysical doubt? His strategy consists in trying to show that Descartes is only concerned with the demands of certainty (understood as a rational confidence in the internal consistency of reason), and does not really care about absolute truth. Now, it seems to me that his main textual evidence in the *Second Replies* (AT- VII- 145) does not really support this dissociation between the search for certainty and the search for truth. The fourth reply taken in its entirety seems to show that the reason why we cannot believe in, or even have the least suspicion of, the absolute falsity of our clear and distinct ideas, is that, once we “became aware that God exists [...] it is *impossible* to imagine that he is a deceiver” (AT- VII- 144, l.16-20). It is this impossibility, and not Descartes’ lack of interest in absolute truth, that “does not allow us to listen to anyone who makes up this kind of story” (AT- VII- 146, l.12-15). I will neither argue for this reading here, nor deal with the other support of Frankfurt’s interpretation, namely, Descartes’ perplexing doctrine of the creation of eternal truths. What I would like to stress is that Frankfurt’s interpretation of Descartes’ overall strategy for defending the consistency of reason, understood as an indirect procedure that removes reasons for doubting truths instead of proving them directly (1989, p. 228), seems to me a perfectly legitimate strategy for defending also the absolute truth of reason. Indeed, once truth is considered as being (or at least as involving) an extrinsic agreement between two distinct terms (thought and object), one of which can only be given to us through its true idea, what could count for us as a direct proof of this correspondence? How could it be possible in this case to give something stronger than an indirect proof? And if this is so, why could we not say that, once we formed the clear and distinct idea of our origin, this idea removes not only any grounds for doubting the internal consistency of reason, but also for doubting the existence of a gap between the rational system of ideas and the ultimate structure of reality? This seems to me exactly the strategy adopted by Spinoza, although, according to his theory of truth, what renders this gap impossible to conceive, namely, the idea’s adequacy is itself one of the two properties constitutive of true ideas.

21 It should be noted that the distinction between an *unjustified affirmation* of the possibility that not-*p* and the exhibition of a *ground for doubting p*, that is, the effective production of an argument supporting the affirmation of the possibility of not-*p*, must be accepted by the pyrrhonic skeptic. In fact, the latter employs the principle according to which for every argument it is always possible to present an opposing, equally strong argument, and he grounds the *epokhé* exactly in this balance of forces. Therefore, in order to accomplish this balance of forces, it is not enough just to affirm the possibility of constructing a counter-argument, but it is necessary to construct it effectively.



**4.8. Objection 2: it is not possible to fulfill the necessary and sufficient condition for the suppression of doubt as long as we ignore God’s existence**

There remains one question about the possibility of forming a clear and distinct idea of God, that is, about the possibility of *fulfilling the necessary and sufficient condition* for the suppression of doubt. The crux of the matter, as Spinoza puts it, consists in determining whether we can form this adequate idea “even though we still doubt whether the author of our nature deceives us in all things” (*PP*, p. 238, *GI/148*). To put it in another way: is it possible to form a clear and distinct knowledge of something while doubting the objective value of this kind of knowledge? Is certainty about the truth of clear and distinct ideas a condition for forming clear and distinct ideas?

Spinoza presents this question as the minor premise of the objection of the circle, which he reformulates and exposes as a syllogism in the end of the prolegomenon to *PP*: “we cannot have a clear and distinct idea of God so long as we do not know whether the author of our nature deceives us” (*GI/149*). What this premise restricts here to the clear and distinct idea of God has evidently a much greater scope than that, as indicated by Spinoza’s response. The premise could be thus reformulated: we cannot have any clear and distinct ideas as long as we ignore whether the author of our nature is deceiving us. So reformulated, this premise evokes a mistake made by Mersenne in the third point of his objections to Descartes’ *Meditations*, where he introduces the problem of the “atheist mathematician”. He writes:

Thirdly, you are not yet certain of the existence of God, and you say that you are not certain of anything, and *cannot know anything clearly and distinctly until you have achieved clear and certain knowledge of the existence of God*. It follows from this that you do not yet clearly and distinctly know that you are a thinking thing (...) Moreover, an atheist is clearly and distinctly aware that the three angles of a triangle are equal to two right angles; but so far is he from supposing the existence of God that he completely denies it. (*AT VII*, 124-125)

Mersenne here heedlessly mistakes the proposition (a): *whoever is not certain of the existence of God cannot be certain of anything*, with the proposition (b): *whoever is not certain of the existence of God cannot know anything clearly and distinctly*. Spinoza, as well as Descartes,<sup>22</sup> accepts (a) but denies (b).

Indeed, it is absurd to turn absolute certainty about the truth of clear and distinct ideas into a precondition for the possibility of their acquisition. In Spinoza’s view, this would amount to affirming that to know, I must first know that I know.<sup>23</sup> If certainty itself was a previous condition for forming clear and distinct ideas, there would never be any metaphysical doubt, since the latter presupposes nothing less than that we have clear and distinct ideas and that we can doubt their truth.

Therefore, the central aspect of the solution of the “Cartesian circle” problem lies in the possibility of forming clear and distinct ideas and rational demonstrations while ignoring our origin and doubting the truth-value of such demonstrations. As Spinoza affirms in *TdIE* §79:

And just as we can arrive at such knowledge of the triangle, even though we may not know whether some supreme deceiver misleads us, so we can arrive at such knowledge of God, even though we may not know whether there is some supreme deceiver. Provided we have that knowledge, it will suffice, as I have said, to remove every doubt that we can have concerning clear and distinct ideas. (*TdIE* §79, *GII/30*)

---

<sup>22</sup> Cf. *Responses to the Second Objections*, *AT- IX-*, p. 111.

<sup>23</sup> Cf. *TdIE* §34, *GII/14-15*.

The difference between the adequate idea of the triangle and the adequate idea of God is that the demonstrative movement of the first leaves untouched the conditions allowing the emergence of the metaphysical doubt, while the second's demonstrative movement, inasmuch as it coincides with the explanation of the contradiction contained in the metaphysical reason for doubting, excludes the only possible reason for doubting our clear and distinct ideas, thus establishing absolute certainty.

To conclude I would like to briefly indicate the intimate connection that exists between Spinoza's solution to the "Cartesian circle" problem and his conception of the self-manifestation of truth (*verum index sui*). Indeed, this solution, based as it is on the logical power of the idea of the most perfect being, offers the most perfect exemplification of that self-manifestation. Spinoza's solution consists in showing that the clear and distinct idea of God logically destroys the metaphysical reason for doubting clear and distinct ideas, and that we can form it while we still doubt the objective value of these ideas. Now, according to Spinoza's theory of truth, clearness and distinctness rest on the intrinsic property that takes part in the very constitution of true ideas, namely, adequacy.<sup>24</sup> Adequacy, understood as the demonstrative process that shows how and why something is necessarily the case, is an integral aspect of what it means for an idea to be true, so that it is not possible for any idea to be true if it lacks this aspect. Yet, to demonstrate something as being necessarily the case is tantamount to excluding the possibility of its negation, and thus the conceivability of its denial. Thus, Spinoza's solution consists in showing that the intrinsic coherence and deductive power of our adequate idea of God, demonstrating its necessary existence and immanence to Nature, establishes the necessary connection linking *adaequatio* and *convenientia*, and renders inconceivable any gap between these two properties.<sup>25</sup> The self-manifestation of God's adequate idea coincides with the self-justification of the possibility of true knowledge, and with the dissolution of the darkness of metaphysical doubt. Spinoza's solution shows hence that his thesis that "truth is the standard both of itself and of false"<sup>26</sup> involves a perfectly consistent conception of the self-justification of the objective value of reason, and that his absolute rationalism is not grounded on an act of blind faith in this value.

## 5. Conclusion

### REFERENCES

- Beysade, J.M. (1997). Sobre o Círculo Cartesiano. *Revista Analytica*, 2(1), 11-36;  
Bolton, M. (1985). Spinoza on Cartesian Doubt. *Nous* 19(3), 379-395;  
Della Rocca, M. (1994). Mental content Content and skepticism Skepticism in Descartes and Spinoza. *Studia Spinozana* 10, 19-421994;

---

<sup>24</sup> See note 3. For a detailed account of Spinoza's concept of adequacy, see Gleizer (2017).[XXX]

<sup>25</sup> Consistent with his denial of any anthropomorphic conception of God, Spinoza refuses to construe God's veracity in terms of goodness or benevolence. Already in his *Metaphysical Thoughts* (Part II, chapter IX), he affirms that: "The same conclusion is also evident from what we said above, viz. that the whole *natura naturata* is only one being. From this it follows that man is part of Nature, which must be coherent with the other parts. Accordingly, it would also follow from the simplicity of God's decree that, if God had created things in another way, he would at the same time have constituted our nature so that we would understand things just as they had been created by God" (GI/267). In the *Ethics*, this necessary coherence among man and Nature, hence, the necessary agreement between our adequate ideas and the structure of reality, will be based on Spinoza's monism, his necessitarianism, the parallelism between the attributes, the status of the human mind as a part of the infinite intellect of God, the identification of our adequate ideas with God's ideas, and the proof that all ideas in God are adequate and agree with their objects. The very clarification of the ontological status of the human mind as a finite mode of the only substance renders thus impossible to conceive it as having any ontological and epistemological isolation from the rest of Nature.

<sup>26</sup> Cf. EHIP43S: "As the light makes both itself and the darkness plain, so truth is the standard both of itself and of the false" (GII/124).

- Descartes, R. (1973). *Oeuvres Philosophiques*. F. Alquié, F. (ed.). Paris: Garnier;
- Descartes, R. (1984). *Philosophical Writings*. Vols. II, ed. J. Cottingham, R. Stoothoff, and D. Murdoch. Cambridge: Cambridge University Press;
- Descartes, R. (1911-1912): *Philosophical Writings*. Vols. I-II, ed. E.S. Haldane and G.R.T. Ross. Cambridge: Cambridge University Press;
- Doney, W. (1975). Spinoza on Philosophical Skepticism. In: Mandelbaum, M. & Freeman, E. (Eds.). *Spinoza: Essays in Interpretation*. La Salle, Illinois: Open Court;
- Frankfurt, H. (1978). Descartes on the Consistency of Reason. In: Hooker, M (Ed.). *Descartes: Critical and Interpretive Essays* (pp. 26-39). Baltimore: Johns Hopkins;
- Frankfurt, H. (1989). *Démons, rêveurs et fous: La defense de la raison dans les Méditations de Descartes*. Épiméthée, Paris: PUF;
- Gleizer, M. (2017). *Vérité et Certitude chez Spinoza*. Paris: Classiques Garnier;
- Hamelin, O. (1984). La théorie de la certitude dans Spinoza, *Les Etudes Philosophiques*, janvier-mars, 97-105, Paris: PUF;
- Hubbeling, H.G. (1967). *Spinoza's Methodology*. Van Gorcum: Assen;
- Spinoza, B. (1924). *Spinoza Opera*. Ed. Carl Gebhardt, 5 vols, Heidelberg: Carl Winters;
- Spinoza, B. (1988: ). *The Collected Works of Spinoza*, vol. 1, . Edited and translated by E. Curley. Princeton: Princeton University Press;
- Spinoza, B. (1951). *Theological-Political Treatise*. In R. H. M. Elwes (ed., tr.);, *The Chief Works of Benedict de Spinoza*. New York: Dover Publications.

---

ALFREDO GATTO

Università Vita-Salute San Raffaele  
gatto.alfredo@hsr.it

---

# DESCARTES E IL PROBLEMA DELLA TEODICEA NELLA PRIMA MODERNITÀ

---

## *abstract*

*The article deals with the critical reception of the Cartesian theory of the eternal truths. It aims to analyze the central role played by Descartes' doctrine in Early Modern Philosophy, with particular regard to the reflections of Leibniz, Malebranche and Spinoza. Indeed, part of their philosophy can be considered as an attempt to criticize the premises of the Cartesian theory in order to avoid their consequences. There are then strong reasons for believing that the importance of the theodicean issues in Early Modern Philosophy is directly related to Descartes' doctrine of the free creation of the eternal truths.*

---

## *keywords*

*Descartes, eternal truths, Theodicy, Leibniz, Malebranche, Spinoza*

La teoria cartesiana sulla libera creazione delle verità eterne rappresenta una delle più radicali riflessioni sull'estensione e le prerogative dell'onnipotenza divina. Sebbene il ruolo svolto dalla dottrina nel pensiero di Descartes sia stato già oggetto di un considerevole numero di studi critici,<sup>1</sup> vi è ancora spazio, a nostro parere, per un'indagine approfondita incentrata sull'importanza da essa giocata nella ricezione del cartesianismo.<sup>2</sup> A questo proposito, crediamo che uno studio esplicitamente dedicato a tale questione possa fornire una nuova chiave di lettura per interpretare il posto occupato da Descartes nel processo di formazione di quella che siamo soliti definire 'filosofia della prima modernità'. Questa teoria ci consente, infatti, soprattutto alla luce della sua ricezione, di tratteggiare un'immagine differente della filosofia cartesiana, una prospettiva che possa mettere in questione l'interpretazione classica di Descartes che abbiamo ereditato dalla ricostruzione storico-filosofica proposta dall'idealismo tedesco.<sup>3</sup> Come è noto, tale lettura venne ripresa in Italia da Gentile<sup>4</sup> e in seguito canonizzata, pur a partire da premesse differenti e con obiettivi opposti, dallo stesso Heidegger, in particolare nei corsi dedicati a Nietzsche<sup>5</sup>.

Secondo questa interpretazione, e al netto delle differenze specifiche presenti nei singoli autori, Descartes è propriamente il filosofo del *cogito*, il primo autore ad aver posto l'auto-riflessione del pensiero al centro della rappresentazione.<sup>6</sup> Descartes sarebbe dunque il legittimo fondatore della filosofia moderna perché ha rioccupato,<sup>7</sup> con la sua nuova nozione di soggettività, il posto tradizionalmente riservato a Dio, facendo della *cogitatio* la condizione di ogni certezza. Sebbene questa lettura si sia trasformata in un luogo comune storiografico,

---

1 Cfr., a semplice titolo di esempio, Agostini, 2004; Alanen, 2008; Curley, 1984; Frankfurt, 1977; Gilson, 2007; Gleizer, 2014; Ishiguro, 1986; Marion, 1981; Rocha, 2016.

2 Benché non manchino degli studi di notevole importanza dedicati proprio alla ricezione della teoria nel dibattito post-cartesiano: cfr., a questo proposito, Easton, 2009; Gasparri, 2007; Marion, 1996, 183-220; Rodis-Lewis, 1981; Schmaltz, 2002.

3 L'interpretazione fornita, in particolare, da Hegel e Schelling è a sua volta mediata dal ruolo decisivo svolto dalla lettura kantiana. Sul rapporto fra Kant e Descartes, si vedano, in particolare, Beyssade, 1996; Ferrari, 1979; Fichant & Marion, 2006; Longuenesse, 2008.

4 Cfr. Gentile, 2003.

5 Cfr. Heidegger, 1950; 1961.

6 Per un'analisi delle differenti interpretazioni del soggetto cartesiano, cfr. Morani, 2007.

7 Sulla legittimità dell'età moderna e sul ruolo svolto da Descartes in questo processo storico, cfr. Blumenberg, 1966-1974.

riteniamo che una simile interpretazione non faccia giustizia alla riflessione cartesiana: Descartes, infatti, non è il pensatore della centralità del *cogito* o della rappresentazione; o meglio: non è *semplicemente* questo. È allora proprio in questo frangente che la sua teoria sulla libera creazione delle verità eterne può contribuire a delineare una prospettiva più articolata del suo pensiero.

Quando parliamo di verità eterne, facciamo tradizionalmente riferimento alle verità matematiche, geometriche, logiche e metafisiche. Fino all'affondo cartesiano, tali verità sono sempre state considerate assolutamente necessarie, costituendo il fondamento del sapere umano. Quello che Descartes afferma, prendendo le distanze da tutta la precedente esperienza medievale, è che queste verità sono state liberamente stabilite e create da Dio – Dio è quindi la loro causa *efficiente*.<sup>8</sup> Se Dio avesse voluto, avrebbe potuto crearne altre, contrarie a quelle ora in vigore, ma ugualmente necessarie.<sup>9</sup> Descartes assegna alle verità eterne uno *status* creaturale: al pari dell'uomo e della natura, tali verità sono creature, il risultato di un *fiat* arbitrario. Queste considerazioni derivano dall'analisi cartesiana dell'estensione dell'onnipotenza divina. In Descartes, Dio non è onnipotente perché può fare tutto ciò che è possibile (questa era la definizione 'medievale' di onnipotenza); al contrario, essendo onnipotente, decide liberamente che cosa rendere possibile.<sup>10</sup> Posta la sua onnipotenza, Dio avrebbe potuto allora creare una differente *ratio entis*, magari incompatibile con quella che regola, attualmente, il sapere umano.<sup>11</sup>

Descartes è convinto inoltre che l'onnipotenza divina, per essere radicalmente pensata, richieda, in particolare, l'*indifferenza* (Dio è indifferente) e l'*incomprensibilità* (Dio è incomprendibile). Tali proprietà, unite alle considerazioni cartesiane sulla natura creata delle verità eterne, generano una frattura e una sproporzione abissale fra l'essenza divina

8 Cfr. Descartes, 2005 (AT, I, pp. 151-152; B Op, n. 32, p. 153): "Mi chiedete in quale genere di causa Dio ha disposto le verità eterne. Vi rispondo che è nello stesso genere di causa nel quale ha creato ogni cosa, cioè come causa efficiente e totale. È certo infatti che egli è autore tanto dell'essenza quanto dell'esistenza delle creature: ora, quest'essenza non è nient'altro che queste verità eterne, che io non concepisco affatto emanare da Dio, come i raggi dal Sole; so, però, che Dio è autore di tutte le cose e che queste verità sono qualche cosa, e di conseguenza che ne è l'autore".

9 Cfr. Descartes, 2005 (AT, IV, p. 18; B Op, n. 454, pp. 1913-1915): "Quanto alla difficoltà di concepire in che modo per Dio è stato libero e indifferente far sì che non fosse vero che i tre angoli di un triangolo fossero uguali a due retti, o in generale che i contraddittori non potessero stare insieme: tale difficoltà si può facilmente superare considerando che la potenza divina non può avere limite alcuno; e poi ancora, considerando che il nostro spirito è finito e creato di una natura tale da poter concepire come possibili le cose che Dio ha voluto fossero veramente possibili, ma non tale da poter anche concepire come possibili quelle cose che Dio avrebbe potuto rendere possibili, ma che ha tuttavia voluto rendere impossibili. La prima considerazione, infatti, ci fa conoscere che Dio non può essere stato determinato a far sì che fosse vero che i contraddittori non possano stare insieme e che, di conseguenza, ha potuto fare l'opposto; l'altra considerazione ci assicura poi che, benché ciò sia vero, noi non dobbiamo cercare di comprenderlo, poiché la nostra natura non ne è capace. E benché Dio abbia voluto che alcune verità fossero necessarie, ciò non equivale a dire che le abbia volute necessariamente: infatti, una cosa è volere che esse fossero necessarie, e tutt'altra volerlo necessariamente, ovvero essere necessitato a volerlo".

10 Cfr. Descartes, 2005 (AT, I, pp. 140-150; B Op, n. 31, p. 151): "Per quanto riguarda le verità eterne, ripeto che sono vere e possibili soltanto perché Dio le conosce come vere e possibili, e non, al contrario, che sono conosciute come vere da Dio quasi fossero vere indipendentemente da lui. E se gli uomini intendessero bene il senso delle loro parole, sarebbero blasfemi qualora dicessero che la verità di qualcosa precede la conoscenza che ne ha Dio, poiché in Dio volere e conoscere non sono che uno; di modo che per ciò stesso che vuole qualcosa, la conosce, e perciò soltanto tale cosa è vera. Non bisogna dunque dire che se Dio non esistesse, queste verità sarebbero comunque vere; l'esistenza di Dio, infatti, è la prima e la più eterna di tutte le verità che possono essere e la sola da cui procedano tutte le altre".

11 Cfr. Descartes, 2005 (AT, V, pp. 223-224; B Op, n. 665, p. 2581): "Ora, a me non sembra che si debba dire di cosa alcuna che non possa essere fatta da Dio. Infatti, poiché tutto ciò che vi è di vero e di bene [*ratio veri et boni*] dipende dalla sua onnipotenza, non oserei neppure dire che Dio non possa far sì che vi sia un monte senza valle, o che uno e due non facciano tre, ma dico solo che Dio mi ha dato una mente tale da non poter concepire un monte senza valle, o una somma di uno e due che non faccia tre, e così via, e che tali cose implicano contraddizione nel mio concetto".

e l'*humana doctrina de Deo*, ossia fra ciò che Dio è e ciò che l'uomo può affermare della sua natura. Poste queste premesse, diviene impossibile qualunque discorso che voglia indagare le ragioni divine, e in particolare le ragioni della creazione. Ciò accade, naturalmente, perché manca un criterio di proporzione che possa essere applicato univocamente all'uomo e a Dio. L'accelerazione nei confronti della tradizione precedente è radicale. A titolo di esempio: in ambito medievale, poiché Dio *non* può creare tutto ciò che implica contraddizione, noi possiamo *sapere*, necessariamente, tutto quello che Dio non poteva e neppure potrà realizzare. Ora, se sappiamo, di necessità, tutto quello che Dio *non* può fare, conosciamo, al netto della sua trascendenza e del divario ontologico che ci separa da Lui, molte cose della sua natura; anzi: un numero infinito di cose, ovvero tutte quelle che, implicando contraddizione, non potevano e non potranno essere oggetto della volontà divina.

Ciò che viene meno con la riflessione cartesiana è proprio questa possibilità di mediazione, che è stata sempre garantita dalla natura immutabile e necessaria delle verità eterne. È per questo che in uno dei passaggi centrali della sua teoria Descartes afferma che, in termini generali, “possiamo essere sicuri che Dio può fare tutto quel che noi possiamo comprendere, ma non che non possa fare quello che non possiamo comprendere, perché sarebbe temerario pensare che la nostra immaginazione abbia la stessa estensione della sua potenza” (Descartes, 2005: AT, I, p. 146; B Op, n. 30, p. 147). Ad esempio, il fatto che l'uomo non possa pensare un mondo che non sia retto dal principio di non contraddizione non implica che Dio non potesse crearlo. Posto il divario che ci separa da Dio, non siamo autorizzati a sostenere che Dio *non* possa realizzare una data cosa *solo* perché ciò appare in contraddizione con la nostra logica *creata*. È precisamente in questo frangente che si insinua una *domanda decisiva*: se non possiamo negare che Dio potesse creare un altro mondo, retto da altri principi logici, siamo legittimati ad escludere che sia tutt'ora nelle condizioni di poterlo fare? Possiamo cioè negare l'estrema eventualità che i principi necessari su cui edificiamo le nostre cattedrali filosofiche possano cambiare? E soprattutto, siamo realmente autorizzati a farlo, posta la natura *incomprensibile e indifferente* di Dio?

Questi interrogativi – è bene specificarlo – non sono stati oggetto della riflessione cartesiana, dato che Descartes non si è mai occupato analiticamente delle implicazioni della sua teoria. Si tratta di una precisazione importante: saranno i suoi critici a sottolineare le possibili conseguenze della dottrina, rilevando alcuni dei potenziali pericoli connessi alla sua formulazione. Non vi sono delle evidenze testuali sull'origine della teoria, dato che non possediamo la lettera di Mersenne che spinse Descartes a prendere posizione all'interno del dibattito. Ad ogni modo, come è stato ampiamente sottolineato dalla storiografia,<sup>12</sup> poiché due delle lettere della primavera del 1630 contengono dei riferimenti ad alcuni passaggi tratti dalle *Disputationes metaphysicae* di Francisco Suárez, è plausibile istituire una correlazione fra l'impostazione difesa dal gesuita spagnolo e l'affondo cartesiano. In questa prospettiva, la dottrina può essere interpretata come una presa di distanza dalla precedente controversia che aveva preso corpo nella Seconda Scolastica.

In termini schematici, il dibattito medievale sullo statuto delle verità eterne può essere analizzato alla luce di due differenti approcci. Secondo il primo, di taglio agostiniano-tomista, le verità eterne, pur essendo in se stesse necessarie ed immutabili, dipendono dall'essenza divina. Se Dio non esistesse, tali verità non sarebbero necessarie, dato che non potrebbero neppure esistere. Ciò, tuttavia, non implica che il loro contenuto dipenda da una decisione divina: poiché Dio non è la loro causa efficiente, tali verità non sono il frutto della sua libera volontà. Esse sono, più propriamente, il contenuto dei pensieri divini, il modello eterno ed

---

<sup>12</sup> Cfr., fra gli altri, gli studi ormai classici di Garin, 1932 e Cronin, 1966.

archetipale di cui Dio si è servito per creare il mondo.<sup>13</sup> Nella cosiddetta Seconda Scolastica, in particolare attraverso la mediazione di autori come Francisco Suárez e Gabriel Vázquez (per citarne solo alcuni), lo statuto delle verità eterne comincia ad acquisire una progressiva autonomia ontologica. Queste verità, oltre a non dipendere dalla causalità efficiente di Dio, sono ora considerate indipendenti dal suo stesso intelletto. Si giunge così all'affermazione radicale secondo cui, anche se Dio non esistesse, le verità eterne non perderebbero affatto la loro assoluta necessità<sup>14</sup>. Alla luce di una simile formulazione, le cose non sono possibili in virtù della relazione che le lega a Dio; al contrario, Dio può pensarle perché già possibili, ossia perché dotate di una natura necessaria che non dipende dalla sua essenza<sup>15</sup>. Seppur implicitamente, la decisione divina di creare il mondo finisce così per essere soggetta ad una sorta di passività originaria, dato che i modelli di cui Dio può servirsi esistono, di necessità, indipendentemente dal suo concorso.

È proprio contro questo secondo approccio che Descartes formula la sua teoria sulla libera creazione delle verità eterne.<sup>16</sup> La dottrina va dunque interpretata attraverso uno sguardo retrospettivo, come una reazione critica nei confronti di un dibattito che aveva già avuto luogo. Possiamo comprendere l'enfasi cartesiana sull'onnipotenza divina solo pensandola attraverso questo preciso contesto storico e teorico. Se poniamo invece l'accento sulla ricezione della teoria, si apre un panorama ben differente. La filosofia post-cartesiana si interrogherà *proprio* sulle conseguenze della dottrina, finendo per rigettarne *in toto* le premesse.<sup>17</sup> I più importanti autori coevi o subito successivi a Descartes forniranno perciò una precisa risposta alle domande poste in precedenza. A loro parere, la radicale indifferenza di Dio, unita all'essenza incomprendibile e arbitraria del suo volere, finisce per gettare un'ombra sulla creazione, privando l'uomo di uno strumento per farsi largo in un mistero che non ha, letteralmente, ragioni, poiché la nozione stessa di *ratio* è stata privata del suo crisma. Il Dio cartesiano, infatti, non agisce in virtù di un ordine di ragioni connaturato alla sua essenza; al contrario, si può parlare di *rationes* solo a partire dall'evento della creazione. Ciò non significa comunque che la *creatio Dei* sia 'irrazionale', perché la supposta irrazionalità divina rinvierebbe, nel caso specifico, ad una razionalità che non appartiene all'essenza di Dio, ma che rappresenta soltanto il risultato di un *fiat* insondabile. Per Descartes, insomma, la creazione, semplicemente, accade, e la logica che ne risulta è una logica, essa stessa, creata, che non può essere utilizzata come un plesso di mediazione per giustificare l'agire divino.

---

13 Cfr., a titolo di esempio, Agostino, 1995 (q. 46, 2, pp. 85-87): "Le idee sono infatti forme primarie o ragioni stabili e immutabili delle cose; non essendo state formate, sono perciò eterne e sempre uguali a se stesse e sono contenute nell'intelligenza divina [...] Ogni cosa è stata dunque creata secondo proprie ragioni. Ma dove crediamo che si trovino queste ragioni ideali se non nella mente stessa del Creatore?"; Tommaso d'Aquino, 1984 (vol. I, q. 10, a. 3, ad 3, p. 214): "Le cose vere e necessarie sono eterne in quanto esistono in un intelletto eterno, che è soltanto l'intelletto divino. Non ne viene perciò che oltre Dio vi sia qualche cosa di eterno"; Tommaso d'Aquino, 1984 (Vol. II, q. 16, a. 7, p. 122): "Se non vi fosse nessuna mente eterna, non vi sarebbe alcuna verità eterna. Ma siccome il solo intelletto di Dio è eterno, soltanto in esso la verità trova la sua eternità".

14 Cfr. Suárez, 1861 (disp. XXXI, s. 12, n. 45, p. 297): "Aequae enim vera est haec conditionalis, si lapis est animal, est sensibilis, ac ista, si homo est animal, est sensibilis; ergo etiam haec propositio, omne animal est sensibilis, per se non pendet ex causa quae possit efficere animal. Unde, si per impossibile nulla esset talis causa, nihilominus illa enuntiatio vera esset, sicut haec est vera, chymaera est chymaera, vel similis".

15 Cfr. Vázquez, 1598 (disp. CIV, c. 3, n. 9, p. 1025): "Res non sunt possibles, quia cognoscuntur, sed ideo cognoscuntur, quia sunt possibles: hoc est, ideo cognoscuntur posse esse, et nulla implicare contradictionem, quia re vera possunt esse".

16 Cfr., a questo proposito, i già citati Garin, 1932 e Marion, 1981.

17 Questa considerazione si applica alla quasi totalità degli autori che si sono confrontati con la teoria. Non mancano, tuttavia, alcune eccezioni, fra cui possiamo ricordare Robert Desgabets, Pierre Cally, Pierre-Sylvain Régis, Louis de la Ville e Pierre Poiret. Sulla posizione di Desgabets, ci permettiamo di rinviare a Gatto, 2017.



Come affermeranno a più riprese, fra gli altri, Leibniz e Malebranche, posto il Dio cartesiano nessuna conoscenza può ritenersi garantita, non essendo possibile ritrovare alcuna necessità nell'indifferenza e arbitrarietà di *quel* Dio. In effetti, al pari del concetto di razionalità, anche la nozione di necessità non rimanda alle decisioni divine, essendo queste il risultato di un atto arbitrario. A tal proposito, i riferimenti non mancano, come mostrano due passaggi tratti, rispettivamente, dai *Saggi di teodicea* di Leibniz e dal decimo *Éclaircissement* di Malebranche:

Se la giustizia è stata stabilita in modo arbitrario e senza alcun motivo, se Dio vi si è imbattuto per una specie di caso, come quando si tira a sorte, la sua bontà e la sua saggezza non vi hanno parte alcuna, anzi non c'è motivo per attribuirgliela. E se è per un decreto del tutto arbitrario, senza alcuna ragione, che ha stabilito o prodotto quel che noi chiamiamo giustizia e bontà, egli le può disfare o mutarne la natura, di modo che non ci sarà alcun motivo di sperare che Dio le osserverà sempre, come invece si può dire che farà, quando si supponga che siano fondate su ragioni (Leibniz, 2005, II, § 177, p. 505).

Se le verità e le leggi eterne dipendessero da Dio, se esse fossero state stabilite dalla libera volontà del Creatore, in una parola, se la Ragione che consultiamo non fosse necessaria e indipendente, mi pare evidente che non vi sarebbe più una scienza vera e propria [...] se l'ordine e le leggi eterne non fossero immutabili in virtù della necessità della loro natura, le più chiare e forti dimostrazioni della religione sarebbero, a mio parere, totalmente distrutte (Malebranche, 1976, p. 132 e 134).<sup>18</sup>

In breve, se l'essenza divina coincidesse con il sistema di attributi delineato da Descartes – vale a dire, se Dio fosse il libero, indifferente e incomprensibile creatore delle verità eterne –, non potremmo escludere che le leggi, eterne e necessarie, che fondano la conoscenza umana possano cambiare, non disponendo di un ordine di ragioni che rifletta la *ratio* immanente alla creazione divina. Per Leibniz e Malebranche, si tratterà allora di rifiutare con forza i presupposti della riflessione cartesiana per recuperare quella relazione con Dio che la teoria di Descartes aveva compromesso.

Ora, prendendo a titolo di esempio il *corpus* di Leibniz, un'analisi delle occorrenze testuali sarebbe, di per sé, sufficiente ad evidenziare l'importanza giocata dalla teoria nello sviluppo e nella formulazione del suo pensiero.<sup>19</sup> Se concentriamo la nostra attenzione sulle opere principali del filosofo tedesco, possiamo subito rilevare come i riferimenti alla dottrina cartesiana abbiano accompagnato buona parte del suo percorso filosofico, a partire dall'epistolario e dalla *Confessio philosophi* fino alla *Monadologia*, passando dal *Discorso di metafisica* e dai *Saggi di teodicea*. Considerazioni simili possono essere fatte per Malebranche e per gli stessi Spinoza e Pascal, e per molti altri autori che svolsero un ruolo decisivo nel dibattito filosofico dell'epoca.

Ad ogni modo, non si tratta semplicemente di portare alla luce la mole di riferimenti alla teoria cartesiana presente nelle opere degli autori citati. Si tratta di sottolineare come il rifiuto delle premesse della dottrina svolga un ruolo *strutturale* nello sviluppo di quelle riflessioni. Sarebbe difficile, infatti, comprendere appieno le considerazioni di tali autori ignorando il confronto-scontro che ha caratterizzato il processo di formazione delle loro prospettive filosofiche. Pertanto – ed è la nostra tesi –, per cogliere la sostanza di questi contributi è necessario

---

<sup>18</sup> Le traduzioni dei passi di Malebranche tratti dall'*Éclaircissement X* sono nostre.

<sup>19</sup> Cfr. Devillairs, 1998.

interpretarli alla luce della dottrina di Descartes, come se essa rappresentasse la condizione – critica – di possibilità delle loro rispettive filosofie. Naturalmente, non è possibile fornire qui un quadro completo di tutti gli elementi a sostegno di questa interpretazione, poiché un esame di questo tipo richiederebbe uno spazio e un lavoro di scavo testuale ben più ampio di quello che possiamo offrire in questa sede. Riteniamo tuttavia opportuno sottolineare, perlomeno, alcuni aspetti principali, così da poter delineare le linee guida generali di una ricerca più ampia.

A tal proposito, possiamo soffermarci sul *problema della teodicea*, che torna a imporsi sulla scena del dibattito filosofico in concomitanza con la diffusione della teoria cartesiana e che rappresenta un aspetto centrale del pensiero di Leibniz e Malebranche.<sup>20</sup> Affinché si ponga la questione della teodicea, intesa ad assolvere Dio al cospetto dello scandalo rappresentato dalla presenza del male nel mondo, è necessario un trittico di attributi, vale a dire l'onnipotenza, la bontà e l'onniscienza di Dio. Ora, il Dio creatore delle verità eterne – un Dio arbitrario, incomprensibile e soprattutto indifferente – esclude alla radice, *agli occhi dei suoi critici*, ogni tentativo di teodicea.<sup>21</sup> D'altronde, se Dio non agisce in vista di *rationes* morali connaturate alla sua essenza, come vuole Leibniz, o seguendo le leggi eterne che la sua saggezza gli impone, come sostiene Malebranche, l'atto della creazione non sarebbe altro che un "decreto assoluto, senza ragione",<sup>22</sup> o privo perlomeno di una *ratio* intrinseca che possa essere indagata dalle sue creature. Insomma, come affermano a più riprese Leibniz e Malebranche, se Dio fosse indifferente, non potremmo considerarlo *necessariamente* un buon Dio<sup>23</sup> – la bontà delle cose, in

---

20 Per uno sguardo sulla teodicea in 'epoca cartesiana', cfr. Brogi, 2006, 58-79, Landucci, 1986 e Nadler, 2008.

21 Nel caso specifico, il corsivo è d'obbligo. Come è noto, infatti, nella quarta meditazione Descartes presenta un tentativo di 'teodicea' volto a giustificare l'errore umano. I critici della teoria, ad ogni modo, non pongono quasi mai in relazione la teodicea della quarta meditazione con quelle che loro ritenevano essere le (possibili) conseguenze della dottrina. E quando ciò accade, sostengono che le ragioni presentate da Descartes nelle *Meditazioni* – in modo particolare, l'evidenza dell'*ego cogito* e il cammino con cui giunge a dimostrare la veracità divina – non siano nelle condizioni di disinnescare le implicazioni legate ai presupposti della teoria. Possiamo riportare ad esempio le osservazioni di due autori 'minori', che si limitavano a dare voce all'*idem sentire* che caratterizzava buona parte del dibattito. Cfr. Daniel, 1690, p. 87: "Car s'il est vraie que la vérité des propositions nécessaires dépend tellement de Dieu, qu'il a pu faire, que celles qui passent pour nécessairement vraies, fussent fausses, il a pu faire que ces deux-ci fussent fausses. *Ce que je conçois clairement est vrai. Être trompeur est une imperfection.* Si Dieu l'a pu faire, qui a dit à Descartes qu'il ne l'a pas fait? Quelle raison a-t-il de le croire, plutôt que le contraire? Dieu le lui a-t-il révélé? Sur son principe, je ne douterai pas seulement en Sceptique, de ces deux propositions, mais j'en douterai sérieusement"; cfr. inoltre Norris, 1701, p. 346: "It is certain that God may absolutely speaking change them [the eternal truths] if he pleases, as depending entirely upon his Will, and how do I know that he has not already done it, or that he will not hereafter do it [?]. Unless you will suppose that he has obliged himself to the contrary. But then this being matter of pure Will and Pleasure too, how shall I know that? It is impossible I should know any of these things except it be by *Revelation*, and even then too, I could not be properly said to know such or such Truths, but only to believe them". Un aspetto ancora più delicato, che non abbiamo però la possibilità di sviluppare in questa sede, concerne il ruolo della dottrina all'interno delle *Meditazioni*. Scribano, 2006, pp. 161-194, ha sostenuto che la teoria non sia presente nel testo in questione (da qui la distinzione tra la metafisica 'pubblica' e 'privata' di Descartes); anzi, sarebbe proprio la separazione tra il contenuto delle *Meditazioni* e le implicazioni della dottrina a giustificare alcuni passaggi dell'opera – si pensi al caso dell'idropico discusso nella sesta meditazione – che non avrebbero potuto trovare altrimenti alcuna giustificazione. A tal proposito, cfr. Scribano, 2000.

22 L'espressione è di Leibniz, ed è tratta da una lettera inviata a Christian Philipp: "Si les choses ne sont bonnes ou mauvaises, que par un effet de la volonté de Dieu, le bien ne sera pas un motif de sa volonté puisqu'il est postérieur à la volonté. Et sa volonté sera un certain décret absolu, sans raison [...] il était donc indifférent à l'égard des choses que nous appelons justes et injustes, et s'il lui avait plu de créer[r] un monde dans lequel les bons fussent malheureux pour toujours et les mauvais (c'est à dire ceux qui ne cherchent qu'à détruire les autres) heureux; cela serait juste. Ainsi nous ne pouvons rien promettre de la justice de Dieu"; Leibniz, 1923–, pp. 787-788 (A, II, 1 B).

23 Descartes aveva infatti affermato che ciò che è buono è tale solo perché Dio ha deciso che fosse così e non altrimenti. La bontà non rappresenta quindi il modello che ha guidato le azioni divine, ma è il risultato della creazione: "A chi presti attenzione all'immensità di Dio, è manifesto che non ci può essere assolutamente nulla che non dipenda da lui;

altri termini, non sarebbe il riflesso della moralità del Creatore, bensì il frutto di una decisione arbitraria; se Dio fosse indifferente, inoltre, non sarebbe legato alla sua creazione, potendo alterare il contenuto di ciò che aveva arbitrariamente decretato; se Dio fosse indifferente, infine, il principio di ragion sufficiente perderebbe ogni funzione, perché non sarebbe più possibile indicare una ragione per le decisioni divine.

È possibile quindi rilevare un rapporto di reciproca e immediata esclusione fra i presupposti della dottrina di Descartes e le riflessioni degli autori citati. Anzi, è forse lecito interpretare tali riflessioni *proprio* come il tentativo di fare fronte alle premesse della teoria con l'obiettivo di negarne le possibili conseguenze. D'altra parte, se l'obiettivo degli autori in questione è giustificare l'operato divino, esponendo le *rationes* della creazione, la teoria cartesiana non solo rende impossibile la soluzione del problema, ma esclude *ab origine* la sua stessa posizione – lo abbiamo già visto: se Dio crea in modo arbitrario e indifferente, e dunque senza 'ragione', non possiamo indagare i motivi per cui ha realizzato ciò che ha realizzato. La posizione della questione della teodicea impone allora il rifiuto delle premesse della dottrina di Descartes. Per raggiungere il loro scopo, i critici di Descartes dovranno quindi, innanzitutto, ritrovare una via di accesso per indagare le ragioni che hanno spinto Dio a creare, recuperando quella relazione fra Creatore e creature che la teoria cartesiana aveva escluso. Il primo passo sarà conferire alle verità eterne lo statuto ontologico che avevano perduto, sottraendole alla contingenza radicale cui erano andate incontro al cospetto del Dio di Descartes. E per due motivazioni principali: da una parte, tale operazione permetterà di attribuire nuovamente ai principi che regolano la conoscenza umana una necessità assoluta, preservando in tal modo l'intrinseca affidabilità del sapere – le verità eterne tornano ad essere gli oggetti *immutabili* dell'intellezione di Dio; dall'altra parte, questo affondo consentirà di sondare l'essenza divina per giustificarne l'operato – l'uomo può così tornare a servirsi delle verità eterne per analizzare, nei limiti del proprio *status*, la *cogitatio Dei*.

A questo proposito, in un passaggio della *Confessio philosophi*, Leibniz si interroga sulla natura delle verità eterne. Il filosofo si riferisce, nello specifico, alle verità della matematica, affermando che il risultato delle sue operazioni non dipende affatto dalla volontà divina. In linea con un approccio tradizionale, modulato secondo il magistero tomista, Leibniz nega che Dio sia la causa efficiente delle verità eterne; al contempo, sostiene però che, senza il suo concorso, esse non potrebbero esistere, dato che sono co-originarie all'apprensione intellettuale della *cogitatio* divina. In breve, le verità eterne non dipendono da una decisione arbitraria di Dio, ma trovano nel suo intelletto la loro condizione di possibilità: esse sono quindi i pensieri che hanno accompagnato il linguaggio della creazione divina, e non sono perciò soggette a cambiare.<sup>24</sup> Ritroviamo considerazioni simili anche in un passo della *Monadologia*: "L'intelletto di Dio è infatti la regione delle verità eterne, ovvero delle Idee da cui tali verità dipendono.

---

non solo nulla di sussistente, ma anche nessun ordine, nessuna legge, o nessuna ragione di vero o di buono [*rationem veri et boni*]: altrimenti, infatti, come si diceva poco sopra, egli non sarebbe stato del tutto indifferente a creare le cose che ha creato. Se una qualche ragione di bene [*ratio boni*] avesse preceduto la sua preordinazione, essa lo avrebbe determinato a fare l'ottimo; ma, al contrario, è perché si è determinato a fare le cose che ora esistono che esse – come è scritto nella *Genesi* – sono molto buone, il che vuol dire che la ragione della loro bontà dipende dal fatto che egli ha voluto farle così"; Descartes, 2009 (AT, VII, pp. 435-436; B Op I, p. 1229).

24 Cfr. Leibniz 2007, pp. 23-25: "Riteniamo forse che vada imputato alla volontà divina se tre per tre fa nove? O pensiamo forse che la diagonale del quadrato è incommensurabile rispetto al lato perché Dio lo ha decretato? [...] Questi teoremi vanno dunque attribuiti interamente alla natura delle cose, ovvero all'idea del nove e del quadrato, e all'intelletto divino nel quale le idee delle cose sussistono eternamente. Il che significa che questi teoremi Dio non li ha posti perché così vuole, ma li ha posti perché pensa, e Dio pensa perché esiste. Infatti, se non vi fosse alcun Dio, tutte le cose sarebbero semplicemente impossibili e il nove o il quadrato seguirebbero il destino comune. Vedi dunque che si danno cose di cui Dio è causa non con la sua volontà ma con la sua stessa esistenza".

Senza l'Intelletto divino, dunque, nessun Reale sarebbe contenuto nel Possibile, e non solo non esisterebbe nulla, ma nulla potrebbe mai esistere" (Leibniz, 2011, p. 77).<sup>25</sup>

Lo stesso ordine di priorità caratterizza l'approccio di Malebranche: nell'*Éclaircissement X*, dedicato al rifiuto della teoria cartesiana, l'oratoriano si sofferma sullo statuto delle verità eterne, criticando con forza la loro natura creata: "quando si pensa all'ordine, alle leggi e alle verità eterne, non si deve cercarne naturalmente la causa, perché non ce l'hanno affatto [...] la natura dei numeri e delle idee intelligibili è immutabile, necessaria, indipendente" (Malebranche, 1976, p. 133). Secondo Malebranche, pensare che la supposta immutabilità di tali verità dipenda da un decreto divino significa allontanarsi dal vero: le verità sono ciò che sono perché oggetto della saggezza di Dio, quella stessa saggezza che lo obbliga ad uniformare le sue azioni al loro contenuto. Pertanto, "le idee intelligibili, o le perfezioni che sono in Dio, che ci rappresentano ciò che è fuori di Dio, sono assolutamente necessarie ed immutabili [...] Le verità sono dunque immutabili e necessarie, al pari delle idee. È sempre stato vero che 2 volte 2 fa 4, ed è impossibile che divenga falso" (Malebranche, 1976, p. 136).

Grazie a queste considerazioni, le verità eterne, immutabili e necessarie, possono nuovamente costituire un prezioso strumento per risalire *in mentem Dei*, consentendo all'uomo di tornare ad indagare la *ratio creationis* con l'intento di giustificarla, così da legittimare, al contempo, l'affidabilità del proprio stesso sapere. Come abbiamo visto, si tratta di un doppio movimento: da un lato, si pongono le condizioni per investigare l'essenza di Dio, così da difenderne l'operato e assolverlo di fronte allo scandalo del male – il recupero dell'intrinseca necessità delle verità eterne è funzionale a questo scopo; dall'altro lato, l'uomo torna in possesso di un ordine necessario di ragioni che gli consente di porre le basi della propria conoscenza, abitando un edificio stabile, i cui fondamenti non possono essere scossi da nessuna volontà o potenza contraria – lo statuto immutabile di tali verità, infatti, esclude che Dio possa intervenire sul loro contenuto, modificandone l'essenza.

Possiamo riformulare quanto detto in questi termini: secondo i suoi avversari, la dottrina di Descartes espone tanto l'universo creato quanto la conoscenza che l'uomo può ottenere di esso ad una contingenza radicale. Il fatto che Descartes avesse difeso l'immutabilità del decreto divino non appare a questi autori una motivazione sufficiente per escludere, alla luce di *questo* Dio, che il contenuto di tali verità, una volta create, non possa nuovamente cambiare. Se non ci sono ragioni sottese alla creazione divina, non ci sono allora neppure ragioni per allontanare lo spettro di una fragilità che potrebbe abitare l'*episteme* umana. Inoltre, l'assenza di una *ratio* incorporata nel mondo creato non offre nemmeno la possibilità di presentare una spiegazione immanente alle decisioni di Dio: sarebbe impossibile, in altri termini, come abbiamo già rilevato, non soltanto fondare una teodicea, ma addirittura porre la questione di un'eventuale giustificazione delle azioni divine.

Per fare fronte a queste due necessità, intimamente legate, è necessario dunque criticare le premesse della teoria cartesiana per disattivarne le possibili conseguenze. Il rifiuto della dottrina di Descartes, pertanto, si pone come una delle pre-condizioni dell'indagine. È per questo che è lecito – e, a nostro parere, imprescindibile – ripensare la speculazione di Leibniz e Malebranche come il controcanto dialettico della tesi cartesiana sulla libera creazione delle verità eterne. Per quanto forte possa apparire quest'ultima considerazione, riteniamo che siano i testi degli autori in esame a giustificarla. Per offrire una maggiore base testuale alle nostre affermazioni, possiamo rileggere alcuni passi centrali del *corpus* leibniziano, a partire dall'*incipit* del *Discorso di metafisica*.

---

25 Cfr. inoltre Leibniz, 2011, p. 79: "Sebbene le verità eterne dipendano da Dio, non bisogna tuttavia credere che esse siano arbitrarie e che dipendano dalla sua Volontà – come pare abbiano sostenuto Cartesio e, dopo di lui, Poiret".

Il primo capitolo del trattato si richiama alla “nozione di Dio più diffusa”, ossia quella che attribuisce al Creatore una “saggezza suprema e infinita” che lo spinge ad agire “nel modo più perfetto, non solo in senso metafisico, ma anche parlando moralmente” (Leibniz, 2000, p. 262). Se sono *queste* le premesse del discorso di Leibniz – vale a dire l’assunzione di un presupposto *metafisico e morale* –, non sorprende che il secondo capitolo si apra proprio con un riferimento critico alle implicazioni della dottrina di Descartes: “Sono dunque molto lontano dalla posizione di chi sostiene che non c’è alcuna regola di bontà e di perfezione nella natura delle cose, o nelle idee che ne ha Dio, e che le opere di Dio non sono buone se non per la seguente ragione formale: che Dio le ha fatte” (Leibniz, 2000, p. 262). Assumendo tale approccio, infatti, prosegue Leibniz, si finisce per distruggere l’amore e la gloria di Dio: “Perché infatti lodarlo per quanto ha fatto, quando sarebbe altrettanto degno di lode se avesse fatto il contrario? Dove finiremmo dunque la sua giustizia e la sua saggezza, se non resta che un certo potere dispotico, se la volontà tiene luogo di ragione [...]?” (Leibniz, 2000, p. 263).

L’ultima considerazione è di particolare importanza, perché costituisce uno dei punti principali del rifiuto leibniziano. La ‘volontà’<sup>26</sup> del Dio cartesiano non si esercita seguendo delle ragioni; di conseguenza, non siamo più nelle condizioni di attribuire al Creatore una giustizia e una saggezza connaturate alla sua essenza. Leibniz è convinto invece che ogni azione della volontà presupponga una ragione determinata, una *ratio* logicamente anteriore all’agire. È per questo che “trovo ben strana l’opinione di certi filosofi, i quali sostengono che le verità eterne della metafisica e della geometria, e di conseguenza anche le regole della bontà, della giustizia e della perfezione, altro non sono che effetti della volontà di Dio, mentre a me pare non siano se non conseguenze del suo intelletto” (Leibniz, 2000, p. 263).

Leibniz rifiuta dunque i presupposti della teoria cartesiana, collocando le verità eterne nella *cogitatio Dei*. È opportuno sottolineare che questo rifiuto ‘apre’ il *Discorso* leibniziano: una volta negata la dottrina, e respinte le sue conseguenze metafisiche e morali, è possibile procedere oltre, edificando quell’edificio epistemico che si sarebbe trovato – *agli occhi di Leibniz*<sup>27</sup> – privo di fondamento, se non si fossero ruscate le premesse della dottrina di Descartes. Nel secondo capitolo del *Discorso* emergono così due elementi: la teoria viene rigettata alla luce delle sue conseguenze morali e in vista delle sue implicazioni metafisiche; inoltre, il suo rifiuto – e ciò vale anche per l’ordine seguito da Leibniz nell’esposizione – è il primo passo che permette al filosofo di avviare la sua indagine, forte di quello strumento di mediazione rappresentato dallo statuto increato delle verità eterne.

Ora, l’idea che la negazione della teoria di Descartes rappresenti una delle pre-condizioni della riflessione leibniziana può trovare conferma anche nella *Monadologia*. Secondo Leibniz, “è la conoscenza delle verità necessarie ed eterne a differenziarci dai semplici animali, e a darci la *ragione* e le scienze, poiché ci eleva alla conoscenza di noi stessi e di Dio” (Leibniz, 2011, p. 71). Il passo è decisivo: la possibilità di conoscere le verità eterne è ciò che ci distingue dagli animali; inoltre, grazie a queste verità, possiamo avere ragione e scienza; infine, attraverso di esse, possiamo elevarci alla conoscenza di quello che siamo e, in modo particolare, possiamo giungere alla conoscenza del Creatore. La funzione ‘mediatrice’ delle verità eterne

---

26 Poniamo questo termine fra virgolette perché il Dio descritto da Descartes non ha, propriamente, una volontà distinta dal suo intelletto: “In Dio, infatti, volere, intendere e creare sono una stessa cosa, senza che l’una preceda l’altra, nemmeno di ragione [*ne quidem ratione*]”; Descartes, 2005 (AT, I, p. 153; B Op, n. 32, p. 153). Pertanto, non si potrebbe parlare di un ‘volontarismo’ cartesiano; in questa sede, utilizziamo il termine seguendo il ragionamento di Leibniz.

27 Anche in questo caso (come nella nota 21), il corsivo è quanto mai opportuno: non stiamo infatti discutendo le conseguenze della teoria in se stesse, ma le implicazioni che gli autori in esame – nel caso specifico, Leibniz – le attribuivano.

si impone ancora una volta, e svolgerà un ruolo fondamentale quando si tratterà di fornire una dimostrazione a priori dell'esistenza di Dio.<sup>28</sup> Non solo: il recupero di un approccio 'tradizionale' alle verità eterne permette a Leibniz di porre a fondamento della conoscenza umana "due grandi principi", il principio di non contraddizione e di ragion sufficiente.<sup>29</sup> Per quanto riguarda il primo principio, Descartes aveva sostenuto che la necessità che gli attribuiamo non fosse *necessariamente* necessaria: il suo attuale valore vincolante dipende dal *fiat* divino, e il fatto che le creature non possano pensare ad un mondo che non sia retto dal *principium firmissimum* non implica che Dio non fosse nelle condizioni di crearlo. Leibniz, lo abbiamo sottolineato, trae dalla premessa cartesiana la seguente conclusione: posta la natura creata delle verità eterne, tutto quello che è stato stabilito da un Dio arbitrario e indifferente – comprese quelle verità che giudichiamo necessarie – potrebbe cambiare; pertanto, anche il principio più saldo di tutti potrebbe vedersi sottrarre il proprio *status* di principio necessario, e senza alcuna ragione. Ora, se il principio di non contraddizione rappresenta una delle due basi della conoscenza umana, per garantire la stabilità del sapere è doveroso rifiutare qualunque approccio ne mini l'intrinseca necessità; se questo è ciò che accade con la dottrina di Descartes, dovremo *allora* rigettarne le premesse.

Oltre ad una ragione metafisica, vi è anche una motivazione morale che spinge Leibniz a porre il rifiuto della teoria cartesiana come punto di partenza dell'indagine. Il principio di ragion sufficiente, che si applica *anche* alle verità contingenti o fattuali, richiede infatti che vi sia una *ratio* implicata nella creazione contingente del mondo. Come è noto, Leibniz riconduce tale ragione alla necessità morale di Dio. Il rapporto di mutua e reciproca esclusione con le premesse cartesiane appare immediatamente: il Dio indifferente di Descartes non si decide alla creazione in virtù di una ragione morale, bensì la stabilisce arbitrariamente, impedendo così alle sue creature di ritrovare nel mondo i modelli della propria *ordinatio*. Anche in questo caso, se il principio di ragion sufficiente è il secondo principio a fondamento dell'umana *episteme*, e se la dottrina cartesiana squalifica *ab ovo* la sua applicazione, per potersene servire sarà necessario rigettare i presupposti della teoria di Descartes. D'altronde, è solo grazie a questo rifiuto che Leibniz può dedicarsi, nell'ultima parte della *Monadologia*, alla descrizione dell'ingegneria sottesa alla creazione, spiegando *perché*, nella sua libertà, Dio abbia creato precisamente ciò che ha creato. L'analisi dei passi di Leibniz potrebbe continuare e applicarsi anche all'opera di Malebranche, tenendo comunque ben ferme le distinzioni tra i due autori. Ad ogni modo, ciò che ci premeva sottolineare in questa sede era l'opportunità di rileggere i due approcci all'interno di un contesto teorico comune, quello aperto dalla dottrina di Descartes e dalla successiva interpretazione delle sue implicazioni. È doveroso precisare, tuttavia, che la soluzione proposta da Leibniz e Malebranche non rappresenta l'unica modalità di affrontare la radicalità della sfida lanciata da Descartes alla modernità. A tal proposito, possiamo ricordare la posizione di Spinoza, che proporrà una soluzione totalmente differente.<sup>30</sup>

La profonda conoscenza che Spinoza aveva della tesi cartesiana è fuori discussione, come dimostrano vari passaggi del suo *corpus*.<sup>31</sup> In questo contributo, concentreremo la nostra

---

28 Cfr. Leibniz, 2011, pp. 77-79.

29 Cfr. Leibniz, 2011, pp. 73-75.

30 La bibliografia sulla relazione fra Spinoza e la dottrina cartesiana è, a nostra conoscenza, limitata. Oltre ai riferimenti presenti nei volumi di Gasparri e Marion indicati nella nota 2, possiamo citare i seguenti contributi: Gatto, 2016; Gleizer, in pubblicazione; Landucci, 1992; Nadler, 2004; Santiago, 2002.

31 Ciò appare chiaramente nei *Principi della Filosofia di Cartesio* e nei *Pensieri Metafisici*. A questo proposito, possiamo citare il titolo del secondo corollario della prop. 12 dei *Principi*, dove è chiara la presupposizione della tesi cartesiana: "Le cose, di per sé, non hanno alcuna essenza che sia causa della conoscenza di Dio, ma, al contrario, Dio è causa anche dell'essenza delle cose". Nella dimostrazione a seguire, Spinoza presenta la posizione di Descartes, che riconduceva l'essenza e l'esistenza delle cose alla causalità efficiente di Dio: "Poiché in Dio non si trova alcuna perfezione che non provenga

attenzione su un passo dell'*Etica*, a sua volta connesso ad un'osservazione già formulata nel *Breve trattato su Dio, l'uomo, e la sua felicità*. Nella prima parte dell'*Etica*, e precisamente nel primo scolio della prop. XXXIII, Spinoza sostiene che le cose non potevano essere prodotte da Dio se non nella maniera in cui sono state prodotte. La distanza che lo separa dall'approccio cartesiano appare quindi in tutta la sua chiarezza. Nello scolio successivo, Spinoza discute una possibile obiezione alla tesi appena presentata che ricorda da vicino le linee guida della dottrina di Descartes:

Ma quelli diranno che nelle cose non c'è né perfezione né imperfezione; e che ciò per cui esse sono dette perfette o imperfette, e buone o cattive, dipende soltanto dalla volontà di Dio; e quindi, se Dio l'avesse voluto, avrebbe potuto far sì che ciò che attualmente è perfezione fosse una grandissima imperfezione, e viceversa «che ciò che attualmente è imperfezione fosse il sommo della perfezione» (Spinoza, 2010/2011, p. 1203: *Etica*, prop. XXXIII, scol. II).

L'obiezione è chiaramente debitrice delle considerazioni legate alla teoria cartesiana. Ciò che è buono o cattivo, perfetto o imperfetto dipende da una volontà arbitraria: se Dio avesse deciso in modo differente, dunque, ciò che ora giudichiamo sommamente imperfetto sarebbe in possesso di tutti i crismi della perfezione. Nelle righe successive, questa posizione è oggetto di una critica serrata.<sup>32</sup> Possiamo tralasciare l'esame delle ragioni spinoziane, non essendo direttamente connesso agli obiettivi del presente contributo. L'aspetto che più ci interessa, infatti, riguarda l'osservazione che chiude lo scolio in esame, dove Spinoza stabilisce una comparazione fra la posizione cartesiana appena riportata ed una riconducibile a quella che sarà la riflessione di Leibniz:

---

da Dio (per P10), le cose per sé non hanno alcuna essenza che possa essere causa della conoscenza di Dio. Ma, al contrario, poiché Dio non ha generato le cose da altro, ma le ha create integralmente (per P12C1), e l'azione della creazione rende nota soltanto la causa efficiente (questa infatti è la definizione di creazione), la quale è Dio, ne segue che le cose prima della creazione erano assolutamente nulla, e quindi Dio è stato causa anche della loro essenza"; Spinoza, 2010/2011, p. 431 (*Principi della Filosofia di Cartesio*, I, prop. XII, corol. II). Un'altra conferma della conoscenza spinoziana della dottrina – e, al tempo stesso, delle modalità con cui poterla incorporare e sviluppare in autonomia – ci è offerta nel nono capitolo della seconda parte dei *Pensieri Metafisici* dedicato alla potenza divina: alcuni "dicono infatti che alcune cose sono possibili per propria natura e non per decreto divino, altre impossibili e infine altre ancora necessarie, e l'onnipotenza di Dio si esercita solo su quelle possibili. Noi, invece, diciamo che Dio è onnipotente dopo aver già dimostrato come ogni cosa dipenda assolutamente dal suo decreto, ma, dopo aver compreso che egli ha decretato determinate cose unicamente in virtù della libertà del suo volere e che inoltre egli è immutabile, asseriamo ora che egli non può in alcun modo agire contro i suoi decreti, e ciò è impossibile per il semplice fatto che contraddice alla perfezione di Dio [...] se gli uomini conoscessero chiaramente l'intero ordine della natura, troverebbero che tutte le cose sono necessarie tanto quanto quelle che si trattano in matematica, ma giacché questo eccede l'umana conoscenza, alcune cose sono giudicate possibili e non necessarie. Perciò o si deve dire che Dio non può nulla, perché tutte le cose sono di per sé necessarie, o che Dio può tutto e la necessità che troviamo nelle cose deriva unicamente dal decreto di Dio"; Spinoza, 2010/2011, p. 591 (*Pensieri Metafisici*, II, cap. IX). Da una parte, il passo citato mostra come Spinoza non solo conoscesse perfettamente la teoria cartesiana, ma sapesse anche collocarla all'interno del dibattito da cui era emersa; dall'altra parte, sottolinea come le premesse della dottrina, che saranno in parte accettate dal filosofo olandese, siano utilizzate per formulare delle considerazioni opposte a quelle presentate da Descartes.

<sup>32</sup> "Ma che altro sarebbe ciò se non affermare apertamente che Dio, il quale concepisce necessariamente ciò che vuole, può, mediante la sua volontà, fare in modo da concepire le cose in un'altra maniera da quella in cui le concepisce? Il che (come dianzi ho mostrato) è una grande assurdità. Io posso, perciò, ritorcere contro di essi il loro argomento nel modo seguente. Tutte le cose dipendono dal potere di Dio. Affinché le cose, dunque, possano essere diverse da quello che sono, anche la volontà di Dio dovrebbe essere necessariamente diversa; ma la volontà di Dio non può essere diversa (come dianzi abbiamo mostrato nel modo più evidente in base alla perfezione di Dio). Dunque neanche le cose possono essere diverse da quello che sono"; Spinoza, 2010/2011, pp. 1203-1205.

Confesso tuttavia che l'opinione che sottomette tutto a una volontà divina, indifferente, e ammette che tutto dipende dal suo beneplacito, s'allontana meno dalla verità che l'opinione di coloro che ammettono che Dio fa tutto in vista del bene. Costoro, infatti, sembra che pongano fuori di Dio qualche cosa che non dipende da Dio, e a cui Dio guarda, come ad un modello, nel suo operare, o a cui egli tende come verso uno scopo determinato. Il che senza dubbio non è altro che sottoporre Dio al fato; cosa, questa, della quale nessuna più assurda si può ammettere intorno a Dio, del quale abbiamo mostrato che è la prima e l'unica causa libera tanto dell'essenza di tutte le cose quanto della loro esistenza. Non c'è dunque ragione che io perda tempo nel confutare quest'assurdità (Spinoza, 2010/2011, p. 1205).

Con questo passo, Spinoza dimostra di aver compreso alla perfezione il rapporto di reciproca esclusione, da noi sottolineato in precedenza, fra la teoria cartesiana e un approccio quale quello leibniziano. Il passo citato ci fornisce inoltre un'informazione ancora più importante: secondo Spinoza, il Dio cartesiano, che agisce in virtù di una volontà libera e indifferente, si allontana meno dal vero di chi sostiene che l'agire divino debba di necessità seguire un ordine morale.<sup>33</sup> Pertanto, benché Spinoza rifiuti le considerazioni cartesiane, è altrettanto convinto che la riflessione del filosofo francese sia di gran lunga preferibile ad un'impostazione che finisce per limitare lo spettro dell'onnipotenza divina, costringendola ad operare sempre *sub conditione*. È interessante notare che questa considerazione era già stata presentata da Spinoza, e quasi negli stessi termini, in un passo del *Breve trattato*, a riprova di quanto tale convinzione fosse radicata nella mente del filosofo:

Il bene è bene solo perché Dio l'ha voluto, e, perciò, Dio può far sì che il male diventi bene. Ma è proprio come se io dicessi che Dio vuole essere Dio e perciò è Dio e, di conseguenza, può non essere Dio, il che è l'assurdità per eccellenza [...] E sebbene coloro i quali dicono che Dio tutto quanto fa, lo fa perché quelle cose sono buone in sé, pensino forse di non differenziarsi da noi, invece differiscono molto: infatti presuppongono un bene che obblighi e vincoli Dio a desiderare che la tale cosa sia buona, la tal altra giusta (Spinoza, 2010/2011, p. 227: *Breve trattato*, IV [6]).

Ben prima dell'*Etica*, Spinoza aveva colto tanto le premesse della teoria sulle verità eterne quanto la perfetta specularità dell'approccio che sarà fatto proprio da Leibniz. Ora, lasciando da parte i meriti di Spinoza come storico della filosofia – cioè l'aver compreso non solo la portata della dottrina cartesiana, ma la centralità del ruolo da essa svolto nel dibattito

---

<sup>33</sup> All'interno di questa prospettiva, non è sorprendente che Leibniz assimili l'arbitrarismo cartesiano e il necessitarismo spinoziano. Secondo il filosofo tedesco, infatti, il *Deus-potentia* di Spinoza – la cui libertà coincide con la propria natura incondizionata – costituisce la radicalizzazione e quasi la 'naturale' prosecuzione delle premesse cartesiane. Se Spinoza "cercava una necessità metafisica negli eventi", convinto che "Dio fosse determinato non dalla propria bontà e dalla propria perfezione [...], bensì dalla necessità della propria natura", Descartes, al contrario, era fautore di un'idea della libertà intensa come "un'indifferenza di equilibrio", Leibniz, 2005 (II, § 174 e 175, pp. 501-503). In entrambi i casi, rileva Leibniz, ciò che approssima i due autori è la convinzione che l'azione *ad extra* di Dio non segua dei modelli che rappresentano le ragioni delle sue scelte. Le considerazioni leibniziane acquisiscono una loro specifica legittimità anche alla luce di un noto passaggio del *Colloquio con Burman*. In effetti, tralasciando la questione relativa all'autenticità dell'opera, possiamo notare come sia stato lo stesso Descartes a istituire una correlazione fra l'indifferenza di Dio e la necessità del suo agire: "Dio, pur essendo indifferente a tutto, ha tuttavia necessariamente decretato così perché necessariamente ha voluto l'ottimo. E qui non si dovrebbe disgiungere nei decreti di Dio la necessità e l'indifferenza: sebbene abbia agito con la massima indifferenza, tuttavia ha al tempo stesso agito in modo massimamente necessario", Descartes, 2009 (AT, V, p. 166; B Op II, p. 1285).



dell'epoca –, possiamo formulare alcune brevi osservazioni sui motivi che hanno spinto il filosofo a considerare la tesi di Descartes più prossima, o meno distante, al suo pensiero. Tale considerazione ci permette di sottolineare come Spinoza, a differenza di Leibniz e Malebranche, non si limiti a criticare i presupposti della teoria, recuperando un approccio riconducibile alla tradizione, ma se ne serva, incorporando parte delle sue premesse. La grande operazione spinoziana consiste nell'accettare integralmente l'idea (propriamente cartesiana) che la causalità efficiente di Dio sia la ragione dell'esistenza e dell'essenza delle cose.<sup>34</sup> Dio non conduce all'esistenza degli enti che riflettono un paradigma eterno ed increato, ma è la causa di quegli stessi modelli. Tuttavia, se Spinoza accetta il presupposto centrale della dottrina, ciò non significa che non lo pieghi a degli obiettivi differenti. Nella formulazione di Descartes, infatti, la teoria era utilizzata per sottrarre alle verità eterne la loro assoluta indipendenza. Essendo liberamente create, tali verità condividevano con il mondo e l'uomo uno statuto comune, facendo parte del medesimo orizzonte *creaturale*. In Spinoza, il primato della causalità efficiente è difeso invece per ricondurre nel dominio ontologico della natura naturata la totalità della creazione, affermando l'assoluta necessità sia delle essenze sia di tutte le esistenze. In altri termini, se in Descartes la comune orizzontalità del creato – il fatto che tutto fosse stato, liberamente, stabilito da Dio, comprese le verità eterne – era funzionale a difendere e preservare con ancora più forza la verticalità della creazione e l'incomprensibilità del Creatore, in Spinoza, al contrario, era affermata per ricondurre ogni cosa nell'ambito di un'immanenza assoluta, *Deus sive natura*. L'immanenza della creazione, insomma, non è sostenuta per preservare un altro dominio ontologico, ma per negarne la legittimità. L'utilizzo che Spinoza fa della teoria contraddice dunque gli obiettivi cartesiani. Mentre per Descartes la natura creata delle verità eterne avrebbe dovuto porre in risalto lo scarto epistemico che divide l'incomprensibilità del Creatore dalle ragioni di cui le creature si servono per comprenderlo, Spinoza rifiuta questa frattura ontologica, eliminando lo iato che Descartes aveva posto a fondamento del suo sistema. La nozione di causalità immanente<sup>35</sup> permette così a Spinoza di risolvere l'essenza divina nello spazio della creazione. In tal modo, come precisa lo stesso filosofo in un passo dell'*Etica*, l'uomo potrà conoscere adeguatamente Dio e l'orizzonte da egli causato (una possibilità, questa, che Descartes non avrebbe mai potuto né accettare né sottoscrivere): “*La mente umana ha una conoscenza adeguata dell'essenza eterna e infinita di Dio*” (Spinoza, 2010/2011, p. 1297: *Etica*, II, prop. XLVII). L'approccio di Spinoza rappresenta un *unicum* nel panorama della ricezione della teoria cartesiana. Il filosofo olandese, infatti, non si limita ad accettarla o rifiutarla, ma sceglie di incorporarne parte delle premesse per giungere a delle conclusioni che finiscono per rovesciare l'ordine di priorità imposto da Descartes. Si può allora parlare di una riformulazione della dottrina: anziché rigettarla a causa delle sue possibili implicazioni, Spinoza decide di assimilarne i presupposti per piegarla ai propri obiettivi metafisici. La scelta di Spinoza lo allontana dal comune atteggiamento assunto nei confronti della teoria da Leibniz e Malebranche e produce delle conseguenze significative all'interno dei rispettivi sistemi filosofici. Se la decisione di rifiutare le premesse della dottrina cartesiana spinge Leibniz e Malebranche a recuperare l'intrinseca necessità delle verità eterne per approssimarsi ad un Creatore che rimane, nella sua essenza, ontologicamente distinto dalle proprie creature,

---

34 Cfr. Spinoza, 2010/2011, p. 1189 (*Etica*, I, prop. XXV): “*Dio è causa efficiente non soltanto dell'esistenza, ma anche dell'essenza delle cose*. Dimostrazione: Se lo neghi, Dio dunque non è causa dell'essenza delle cose; e quindi (per A4) l'essenza delle cose può essere concepita senza di Dio; ma ciò (per P15) è assurdo. Dio, dunque, è causa anche dell'essenza delle cose”.

35 Cfr., ad esempio, Spinoza, 2010/2011, p. 1189 (*Etica*, I, prop. XVIII): “*Dio è causa immanente, e non transitiva, di tutte le cose*”.

Spinoza elimina alla radice la possibilità di questo scarto: così facendo, pone le condizioni per conoscere adeguatamente l'essenza eterna ed infinita di Dio, rimuovendo le ragioni stesse che avevano spinto gli altri protagonisti a presentare il loro progetto di teodicea.

Le considerazioni svolte, che dovranno certamente essere oggetto di un studio più approfondito, ci permettono forse di ampliare le categorie ermeneutiche con cui interpretare la filosofia moderna, perlomeno nell'atto della sua fondazione. È infatti riduttivo continuare a rileggere quest'epoca storica *soltanto* alla luce di un semplice passaggio di consegne, quasi che la centralità assegnata al soggetto rappresentasse la ri-occupazione di un registro teorico riservato tradizionalmente a Dio.<sup>36</sup> Dovrebbe essere ormai chiaro che le questioni poste dall'onnipotenza divina, soprattutto a partire dalla loro riformulazione cartesiana, continuano ad occupare la scena del dibattito filosofico, spingendo i più autorevoli rappresentanti a prendere posizione. Le varie strategie presentate e le differenti risposte che sono state fornite possono, tuttavia, essere interpretate a partire da un'esigenza comune, possono cioè essere lette come il tentativo di affrontare le implicazioni della dottrina di Descartes, in particolare il radicale contingentismo della conoscenza umana che potrebbe derivarne. Per queste ragioni, il filtro offerto dalla teodicea, strettamente legato alla tesi cartesiana, offre un'ulteriore chiave di lettura per interpretare la prima modernità. Come abbiamo cercato di sottolineare nel corso dell'articolo, si tratta comunque di un tentativo di teodicea oltremodo paradossale: non si cercherà solo di giustificare e assolvere l'operato divino, ma di proteggere la stabilità del sapere umano al cospetto del Dio onnipotente cartesiano.

#### BIBLIOGRAFIA

- Agostini, I. (2004). "Nihili nullae proprietates esse possunt" (AT V 223). Impossibilità del vuoto, assioma del nulla e verità eterne in Descartes. In A. Del Prete (Ed.), *Il Seicento e Descartes. Dibattiti cartesiani* (pp. 3-24). Firenze: Le Monnier;
- Agostino d'Ippona (1995). Ottantatré questioni diverse. In Agostino d'Ippona, *La vera religione*. Ed. G. Ceriotti. Roma: Città Nuova;
- Alanen, L. (2008). Omnipotence, Modality and Conceivability. In J. Broughton & J. Carriero (Eds.), *A Companion to Descartes* (pp. 353-371). Malden: Blackwell;
- Arbib, D. (2017). *Descartes, la métaphysique et l'infini*. Paris: Puf;
- Beysade, J.-M. (1996). La critique kantienne du "cogito" de Descartes (sur le paragraphe 25 de la Déduction transcendantale). In C. Ramond (Éd.), *Kant et la pensée moderne: alternatives critiques* (pp. 47-62). Bordeaux: Presses Universitaires de Bordeaux;
- Blumenberg, H. (1966-1974). *Die Legitimität der Neuzeit*. Frankfurt am Main: Suhrkamp Verlag;
- Brogi, S. (2006). *I filosofi e il male. Storia della teodicea da Platone ad Auschwitz*. Milano: Franco Angeli;
- Cronin, T. J. (1966). *Objective Being in Descartes and Suárez*. Rome: Gregorian University Press;
- Curley, E. (1984). Descartes on the Creation of the Eternal Truths. *The Philosophical Review*, 93, 569-597;
- Daniel, G. (1690). *Voyage du Monde de M. Descartes*. Paris: chez Veuve de S. Bénard.
- Descartes, R. (2005). *Tutte le lettere. 1619-1650* (B Op). Ed. G. Belgioioso, con la coll. di I. Agostini, F. Marrone, F. A. Meschini, M. Savini e J.-R. Armogathe. Milano: Bompiani;
- Descartes, R. (2009). *Opere: 1637-1649* (B Op I). Ed. G. Belgioioso, con la coll. di I. Agostini, F. Marrone, M. Savini. Milano: Bompiani;
- Descartes, R. (2009). *Opere postume. 1650-2009* (B Op II). Ed. G. Belgioioso, con la coll. di I. Agostini, F. Marrone, M. Savini. Milano: Bompiani;

---

<sup>36</sup> A questo proposito, si veda il fondamentale contributo di Marion, 1996, pp. 3-47. Sulla 'finitezza' del cogito, cfr. inoltre Arbib, 2017 e Olivo, 2005.

- Devilleirs, L. (1998). *Descartes, Leibniz. Les vérités éternelles*. Paris: Puf;
- Easton, P. (2009). What is at Stake in the Cartesian Debates on the Eternal Truths?. *Philosophy Compass*, 4, 2, 348-362;
- Ferrari, J. (1979). *Les sources françaises de la philosophie de Kant*. Paris: Klincksieck;
- Fichant, M. & Marion, J.-L. (2006). *Descartes en Kant*. Paris: Puf;
- Frankfurt, H. (1977). Descartes on the Creation of the Eternal Truths. *The Philosophical Review*, 85, 36-57;
- Garin, P. (1932). *Thèses cartésiennes et thèses thomistes*. Paris: Desclée de Brouwer;
- Gasparri, G. (2007). *Le grand paradoxe de M. Descartes. La teoria cartesiana delle verità eterne nell'Europa del XVII secolo*. Roma: Leo S. Olschki;
- Gatto, A. (2016). A teoria cartesiana das verdades eternas na interpretação de Espinosa. *Cadernos Espinosanos*, 35, 269-294;
- Gatto, A. (2017). La puissance épuisée. Robert Desgabets et les vérités éternelles. *Recherches Philosophiques*, 4, 129-152;
- Gentile, G. (2003). *Sistema di logica come teoria del conoscere* (2 vols.). Firenze: Le Lettere;
- Gilson, É. (2007). *La liberté chez Descartes et la théologie*. Paris: Vrin (1<sup>re</sup> éd., Paris: Félix Alcan, 1913);
- Gleizer, M. A. (2014). Descartes e a livre criação das verdades eternas. In M. A. Gleizer, *Metafísica e conhecimento. Ensaio sobre Descartes e Espinosa* (pp. 133-156). Rio de Janeiro: EDUERJ;
- Gleizer, M. A. (in pubblicazione). Spinoza et la théorie cartésienne de la libre création des vérités éternelles. In *La générosité à l'œuvre. Hommage à Jean-Marie Beysade*. Paris: Classiques Garnier;
- Heidegger, M. (1950). Die Zeit des Weltbildes. In M. Heidegger, *Holzwege*. Frankfurt am Main: Verlag Klostermann;
- Heidegger, M. (1961). *Nietzsche*. Pfullingen: Verlag Günther Neske;
- Ishiguro, H. (1986). The status of Necessity and Impossibility in Descartes. In A. O. Rorty (Ed.), *Essays on Descartes's Meditations* (459-472). Berkeley: University of California Press;
- Landucci, S. (1986). *La teodicea nell'età cartesiana*. Napoli: Bibliopolis;
- Landucci, S. (1992). Sulle "verità eterne" in Spinoza. In D. Bostrenghi (Ed.), *Hobbes e Spinoza. Scienza e politica* (pp. 23-51), Atti del Convegno Internazionale Urbino, 14-17 ottobre. Napoli: Bibliopolis;
- Leibniz, G. W. (1923-). *Sämtliche Schriften und Briefe*. Darmstadt-Leipzig-Berlin: Akademie der Wissenschaften (A);
- Leibniz, G. W. (2000). *Discorso di metafisica*. In G. W. Leibniz, *Scritti filosofici*. Eds. M. Mugnai & E. Pasini. Torino: Utet;
- Leibniz, G. W. (2005). *Saggi di Teodicea sulla bontà di Dio, la libertà dell'uomo e l'origine del male*. Ed. S. Cariati. Milano: Bompiani;
- Leibniz, G. W. (2007). *Professione di fede del filosofo*. In G. W. Leibniz, *Dialoghi filosofici e scientifici*. Ed. F. Piro. Milano: Bompiani;
- Leibniz, G. W. (2011). *Principi della filosofia o Monadologia*. Ed. S. Cariati. Milano: Bompiani;
- Longuenesse, B. (2008). Kant's 'I think' versus Descartes' 'I am a Thing That Thinks'. In D. Garber & B. Longuenesse (Eds.), *Kant and the Early Moderns* (pp. 9-31). Princeton: Princeton University Press;
- Malebranche, N. (1976). *Éclaircissements*. Éd. G. Rodis-Lewis. In *Œuvres complètes*, 20 vols., dir. A. Robinet. Paris: Vrin;
- Marion, J.-L. (1981). *Sur la théologie blanche de Descartes. Analogies, création des vérités éternelles et fondement*. Paris: Puf;
- Marion, J.-L. (1996). L'altérité originaire de l'ego. In J.-L. Marion, *Questions cartésiennes II. Sur l'ego et sur Dieu* (3-47). Paris: Puf;

- Marion, J.-L. (1996). Création des vérités éternelles. Principe de raison. Malebranche, Spinoza, Leibniz. In J.-L. Marion, *Questions cartésiennes II. Sur l'ego et sur Dieu* (pp. 183-220). Paris: Puf;
- Morani, R. (2007). *Soggetto e modernità. Hegel, Nietzsche, Heidegger interpreti di Cartesio*. Milano: Franco Angeli;
- Nadler, S. (2004). Les vérités éternelles et l'autre monde: les racines juives de Spinoza. *Les Études philosophiques*, 4, 71, 507-522;
- Nadler, S. (2008). *The Best of All Possible Words. A Story of Philosophers, God, and Evil in the Age of Reason*. New York: Princeton University Press;
- Norris, J. (1701). *An Essay Towards the Theory of the Ideal or Intelligible World*. London: S. Manship;
- Olivo, G. (2005). *Descartes et l'essence de la vérité*. Paris: Puf;
- Rocha, E. M. (2016). *Indiferença de Deus e o mundo dos humanos segundo Descartes*. Curitiba: Kotter Editorial;
- Rodis-Lewis, G. (1981). Polémiques sur la création des possibles et sur l'impossible dans l'école cartésienne. *Studia Cartesiana*, 2, 105-123;
- Santiago, H. S. (2002). Descartes, Espinosa e a necessidade das verdades eternas. *Cadernos de História e Filosofia da Ciência* (Campinas-UNICAMP), 12, n. 1-2, 315-325;
- Schmaltz, T. (2002). *Radical Cartesianism. The French Reception of Descartes*. Cambridge: Cambridge University Press;
- Scribano, E. (2000). Quel che Dio non può fare. Descartes e i limiti della potenza divina. In G. Canziani & M. A. Granada & Y. Ch. Zarka (Eds.), *Potentia Dei. L'onnipotenza divina nel pensiero dei secoli XVI e XVII* (pp. 335-350). Milano: Franco Angeli;
- Scribano, E. (2006). *Angeli e beati. Modelli di conoscenza da Tommaso a Spinoza*. Roma-Bari: Laterza;
- Spinoza, B. (2010/2011). *Tutte le opere*. Ed. A. Sangiacomo. Milano: Bompiani;
- Suárez, F. (1861). *Disputationes Metaphysicae*. In F. Suárez, *Opera Omnia*, vols. 25-26. Ed. Nova par C. Berton. Paris: Ludovico Vivès;
- Tommaso d'Aquino (1984). *La Somma Teologica*. Ed. P. A. Balducci. Bologna: Edizioni Studio Domenicano;
- Vázquez, G. (1598). *Commentarium ac disputationum in primam partem S. Thomae*. Alcalà: apud Juan Gracián.



---

RAFFAELE ARIANO

Università Vita-Salute San Raffaele  
raffaele.ariano1985@gmail.com

---

# GEOGRAFIA E FILOSOFIA. RIFLESSIONI SU *PENSIERO VIVENTE* DI ROBERTO ESPOSITO A PARTIRE DA SPINOZA, CAVELL E FOUCAULT

---

## *abstract*

*The article focuses on the so-called 'Italian theory', with special reference to its assessment in Roberto Esposito's 2010 book Living Thought. The article puts forth three main arguments. First, there is an analogy between the 'geo-philosophical' strategy chosen by Esposito to address Italian thought and the strategy pursued by Stanley Cavell when he reflects on Emerson and the specificity of American philosophy. Second, the main traits attributed by Esposito to Italian thought show significant parallels with key concepts in Michel Foucault's philosophy. Third, interpretations of Spinoza by authors such as Gilles Deleuze and Toni Negri have been instrumental for the flourishing of the philosophical debate that prepared Esposito's elaboration of a canon of Italian thinkers. The article thus hints that Italian theory could be an interesting case study for the reflection on the relations between geography and philosophy.*

---

## *keywords*

*Spinoza, Esposito, Foucault, Italian theory, Cavell, the order of things*

Nel corso delle pagine seguenti cercherò di avanzare delle considerazioni – che mi permetterei di definire ‘ironiche’ per motivi che spero diventeranno chiari nelle battute conclusive del mio intervento – sui rapporti tra filosofia e geografia, e altrettanto su quelli tra tradizioni interpretative e innovazione filosofica.

Il tema che ho scelto – e cioè il rapporto tra Spinoza e quel complesso di pensatori e filosofie che sono stati rubricati in tempi recenti sotto l’etichetta di *Italian Theory*, o anche di *Italian Thought* (vedi Claverini, 2016; 2017) – ha in se stesso qualcosa d’imbarazzante. Non si tratta solo del fatto che il sottoscritto non è personalmente uno specialista né di Spinoza, né di *Italian Theory* – e lo dico quindi da subito: questo mio intervento vuol essere un contributo, tra virgolette, “di servizio”, utile a raccordare i diversi corni teorici di questa interessante e complessa giornata di studi. Non si tratta nemmeno del fatto che forse affronterò Spinoza in modo più rapido di quanto il titolo dell’intervento potrebbe indurre ad aspettarsi. È imbarazzante anche per un motivo più sostanziale. La rivendicazione di una specificità non solo geografica, ma teorica, ovvero genuinamente filosofica, del pensiero italiano, è stata tale da richiedere una precisa fatica ermeneutica e teorica, un lavoro di scavo genealogico e un posizionamento teorico escogitati ad arte, i quali, a quanto sembra, arrivano solo tardivamente e a prezzo di grandi resistenze. Se tutto questo è vero, la mia scelta di contribuire alla riflessione sulla specificità di quella corrente del pensiero meridiano che è la filosofia italiana proprio col riferimento a un filosofo, che italiano non è, potrà sembrare ad alcuni un modo per eludere la questione.

In realtà, una delle prime scoperte che si fanno addentrandosi nel problema dei rapporti tra geografia e filosofia, tra territorio e filosofia, è che la rivendicazione della sostanzialità di questi rapporti è soggetta alle stesse complessità e agli stessi andirivieni che caratterizzano i rapporti tra, per esempio, letteratura e geografia, o tra arti figurative e geografia. Scopriamo così ancora una volta che, a dispetto della sua capacità totalizzante di parlare del mondo intero e quindi di contenerlo, anche la filosofia è contenuta nel mondo e, segnatamente, in quella porzione di mondo che sono le culture umane. (Con ciò, tra l’altro, spero di riuscire a ricollegarmi idealmente anche al tema generale delle tre giornate di convegno, che recita appunto: “*I metodi della filosofia*”. Giacché tra i metodi della filosofia vi è appunto anche quello storico-ermeneutico variamente declinato, e declinabile appunto, come fanno a vario titolo alcuni degli studiosi che hanno contribuito al convegno, come storia critica delle idee). Anche le filosofie, insomma, rivelano non soltanto un radicamento storico e geografico alla

stregua di tutte le altre forme della cultura umana, ma lo fanno negli stessi modi impuri e complessi, in modi caratterizzati dagli stessi corsi e ricorsi e dagli stessi andirivieni dialettici di dentro e fuori, di territorializzazione e deterritorializzazione – per dirla col vocabolario deleuzeggiante utilizzato da Roberto Esposito (2010) nel suo *Pensiero vivente*, che prendo qui come testo di riferimento della riflessione sull'*Italian Theory*. Non dovremo stupirci, perciò, di scoprire che, nell'edificazione di un canone della filosofia italiana, debbano intervenire così prepotentemente filosofi non italiani – tra di loro, Spinoza – né che il ritrovamento di una specificità, di una “origine e attualità della filosofia italiana”, per dirla col sottotitolo del libro di Esposito, abbia un carattere così spiccatamente programmatico e retrospettivo.

Nel suo libro sull'*Italian thought*, Esposito sostiene appunto che una peculiarità geofilosofica del pensiero italiano sia quella di innescare una costante dialettica tra territorializzazione e deterritorializzazione. Sin dalla grande stagione dell'umanesimo, la filosofia italiana è nata e cresciuta in un contesto di forte decentramento e frammentazione politica, un contesto caratterizzato socio-politicamente dall'esperienza dei comuni e delle signorie. Suo tratto tipico, spiega Esposito, era inoltre un alto tasso di emigrazione intellettuale. Da Firenze, Venezia, Milano, Roma e Napoli intellettuali e artisti italiani emigravano in tutt'Europa, cagionando fenomeni come le precoci traduzioni francesi delle opere di Dante, Petrarca e Boccaccio, la diffusione del platonismo fiorentino di Ficino e Pico, la fortuna internazionale del pensiero politico di Machiavelli. Insomma, il grande pensiero filosofico italiano è un pensiero cresciuto in assenza di uno stato-nazione ed è forse anche per questo – se Esposito ha ragione, e una qualche continuità può essere tracciata che connetta in un'unica genealogia Giordano Bruno e Vico con pensatori come Agamben, Toni Negri e lo stesso Esposito – , è forse per quest'assenza di un riferimento allo stato-nazione che i pensatori italiani si sono rivelati e si stanno rivelando così cruciali per il dibattito filosofico internazionale dagli anni Novanta ad oggi, in un'epoca in cui la rilevanza dello stato-nazione si è, secondo taluni, così fortemente ridimensionata.

Queste dinamiche di nomadismo intellettuale tornano insomma anche in tempi recenti. Risultano ben chiare quando si consideri che il dibattito recente sulla specificità della filosofia italiana è stato sì condotto prevalentemente da studiosi italiani, ma in un contesto accademico ed editoriale spesso anglofono. La discussione confluita nel testo del 2010 di Esposito comincia probabilmente nel 1996 con l'antologia *Radical Thought in Italy: A Potential Politics*, curata dal coautore di *Impero* Michael Hardt (Hardt & Negri, 2003) e da Paolo Virno, egli stesso uno dei più letti esponenti del pensiero radicale italiano. Se escludiamo dal dibattito il breve pamphlet polemico di Toni Negri pubblicato nel 2005, intitolato *La differenza italiana*, possiamo citare la raccolta del *The Italian Difference: Between Nihilism and Biopolitics* (Chiesa & Toscano, 2009), ancora una volta in lingua inglese. Non è un caso, quindi, che Dario Gentili, tornando a riflettere sulla questione nel suo *Italian Theory. Dall'operaismo alla biopolitica* (2012), abbia scelto di mantenere l'inglese sin nel titolo del suo libro. “*Italian Theory*” è insomma un'etichetta filosofico-storiografica che inquadra il pensiero italiano da fuori, e segnatamente dalla prospettiva del mondo accademico anglo-americano. Esposito traccia un'analogia tra quest'adozione del radicalismo filosofico italiano da parte dell'accademia americana e analoghe adozioni avvenute nel '900: nei primi decenni del secolo, quella del pensiero filosofico d'area tedesca – dal positivismo logico alla Scuola di Francoforte; in epoche più recenti, quella del post-strutturalismo e del decostruzionismo, i quali diedero origine, una volta giunti nel mondo anglosassone, all'etichetta storiografica di *French Theory*, sulla quale quella di *Italian Theory* è ovviamente modellata (Esposito, 2010, p. 3).

C'è inoltre un secondo aspetto di quest'apertura all'altro da sé del pensiero italiano, consistente nel fatto che i recenti esponenti dell'*Italian Theory*, pur mantenendo secondo Esposito una forte specificità e una connessione con la lunga tradizione di pensiero italiano da



lui identificata, desumono nondimeno molti dei loro temi e dei loro concetti dal precedente dibattito filosofico europeo, da quello francese in particolare. La forte influenza di pensatori come Foucault e Deleuze sul pensiero di Agamben, Esposito e Negri è cosa nota. Altrettanto nota è la centralità per questi pensatori italiani della categoria di biopolitica.

Si segnala qui una tangenza coi miei recenti interessi di ricerca che mi permettono di affrontare aprendo una parentesi, sperando che il *detour* non risulti troppo estraneo al discorso. Mi è necessario fare un salto geografico non di poco conto. L'operazione di Esposito, infatti, presenta qualche analogia col tentativo condotto da Stanley Cavell di rinvenire una specificità americana del filosofare (1981, 1989, 1990). Cavell non si considera semplicemente un filosofo a cui accade di vivere in America, ma un filosofo americano, un filosofo per cui è cruciale interrogarsi su cosa significhi fare filosofia in America. Da ciò si può intuire un movimento critico nei confronti dell'egemonia, nelle università americane, della filosofia analitica, un movimento critico che in modi diversi potrà essere rinvenuto anche nei successivi processi d'assorbimento nel mondo accademico americano della *French Theory*, prima, e dell'*Italian Theory*, poi. Ma l'analogia non si ferma qui. Il punto è che Cavell scrive come se la distinzione tra filosofia analitica e continentale non esistesse affatto. Cavell si forma filosoficamente sulla filosofia del linguaggio ordinario di Wittgenstein e Austin, ma scrive in seguito anche su Nietzsche, su Heidegger, su Lévinas. Per converso, questa interpretazione non analitica di Wittgenstein, così come l'utilizzo di Heidegger e Nietzsche gli permettono di tornare con sguardo nuovo al pensiero americano, in cerca di una genealogia diversa. Cavell scopre così che il pensiero filosofico americano non inizia, come si crede abitualmente, col pragmatismo, bensì con autori che tradizionalmente non venivano rubricati nemmeno come filosofi, ovvero con Emerson e Thoreau. Per dirla in estrema sintesi: l'americano Cavell fuoriesce dalla filosofia americana istituzionale recuperando filosofi europei che ne erano esclusi e può per questo volgere nuovamente lo sguardo all'America, riabilitando come filosofi pensatori che filosofi non erano considerati e, così facendo, spostando i confini stessi dell'argomentare e della scrittura filosofica negli Stati Uniti. Recuperando alla filosofia letterati come Emerson e Thoreau, Cavell pone insomma al centro della riflessione filosofica il tema dei rapporti tra stile e pensiero, tra letteratura e filosofia.

Esposito fa, per certi versi, qualcosa di simile. Ci dice che negli ultimi decenni alcuni pensatori italiani hanno reinterpretato in modo originale il pensiero post-strutturalista francese. Così facendo hanno posto al centro del loro pensiero le categorie di vita, storia e politica, delle quali l'utilizzo che Esposito fa del concetto foucaultiano di biopolitica è appunto una sintesi. Così facendo, si è dischiusa per loro la possibilità di tornare indietro e costruire una nuova genealogia del pensiero italiano che vada da Machiavelli fino a noi, una genealogia nella quale il concetto di vita assume un ruolo centrale e in cui vengono trattati come filosofi anche pensatori che abitualmente venivano esclusi dal canone ufficiale della filosofia. Se Vico, Gramsci, Croce e Gentile sono certamente filosofi con la "F" maiuscola, più sfumata è l'appartenenza al canone filosofico maggiore di Machiavelli e Giordano Bruno, o ancora, per approdare a tempi più recenti, quella di un Mario Tronti. Ancor più evidente è il discorso se si considera che in questa genealogia della filosofia italiana Esposito include anche storici, poeti e critici letterari. Al *Saggio storico sulla rivoluzione napoletana* di Vincenzo Cuoco Esposito dedica alcune delle pagine più ispirate del suo testo (2010, pp. 99-111). Altrettanto importante, per chiudere qui i miei esempi, è l'inclusione in *Pensiero vivente* di figure come Giacomo Leopardi, Francesco De Sanctis e Pier Paolo Pasolini. La questione dello stile viene, tra l'altro, trattata in modo esplicito anche da Esposito: non è infatti un caso, a suo avviso, se a filosofare in Italia sono stati spesso anche i non filosofi di professione. È l'intrinseca apertura del pensiero Italiano alla vita, alla storia, all'orizzonte complesso del conflitto – in una parola l'apertura della filosofia italiana a quello che Esposito chiama il "non filosofico" (p. 12) – a farle porre

naturalmente grande attenzione allo stile, alla molteplicità e complessità dei generi espressivi, all'interscambio reciproco tra le discipline. Scrive Esposito in un passo su Vico: "La stessa modalità – concreta, corporea, materica – del linguaggio adoperato non è che il modo di tradurre nel presente della conoscenza quell'ampia zona di esperienza preriflessiva da cui essa proviene e da cui non può mai completamente distaccarsi, se non vuole chiudersi in un gergo astratto e prosciugato" (2010, p. 72). Il linguaggio della filosofia italiana ha una modalità concreta, corporea, materica. La filosofia italiana ha uno stile, o meglio ancora – più stili. Trovo interessante che tanto per Cavell quanto per Esposito la ricerca di una specificità territoriale in filosofia includa un movimento anti-disciplinare nei confronti della sussunzione della filosofia *tout court* all'interno della concezione analitica della filosofia e ancor più che, per entrambi, tale operazione passi attraverso la riabilitazione filosofica dei non-filosofi e la valorizzazione dello stile in filosofia, della filosofia come scrittura dotata di uno stile.

Ad ogni modo, l'analogia tra Esposito e Cavell, strana forse sin dappprincipio, si ferma qui, dal momento che, mentre quello di Cavell è, in senso lato, pur sempre un pensiero della svolta linguistica, la peculiarità del pensiero italiano è secondo Esposito costituita proprio dalla sua capacità di uscire dall'*impasse* che tutte le filosofie della svolta linguistica si starebbero trovando ad attraversare. Al posto del linguaggio, come dicevo, il pensiero italiano ruota a suo avviso attorno alla categoria di vita. È qui che iniziamo ad avvicinarci a Spinoza. Esposito menziona Spinoza in alcuni dei luoghi-cardine del suo libro. Traccia un parallelo tra Spinoza e Machiavelli (2010, p. 55-56), lo lascia aleggiare (e non avrebbe potuto essere altrimenti) sulle pagine dedicate a Giordano Bruno (pp. 60-71), lo fa tornare inevitabilmente nel paragrafo incentrato sulla riflessione di Toni Negri (pp. 218-224). Esposito arriva a descrivere Spinoza – testuali parole – come "il più 'italiano' dei filosofi moderni" (p. 32).

Esposito nel suo libro non approfondisce direttamente il pensiero di Spinoza. Per capire queste affermazioni sull'italianità del filosofo olandese, perciò, è necessario riflettere sulla caratterizzazione generale del pensiero italiano da lui proposta. È qui che mi accingo ad avanzare l'unica vera e propria tesi di questo mio intervento. Innanzitutto, sono convinto che i tre plessi teorici portanti del pensiero italiano identificati da Esposito possano essere letti ciascuno come la riformulazione e lo spostamento retrospettivo in un contesto italiano di uno specifico aspetto della proposta teorica di Michel Foucault. Vi è inoltre una seconda cosa da rilevare. Esposito afferma in più circostanze che la filosofia italiana – e questa volta stiamo facendo riferimento non a quella tradizionale, bensì a quella degli ultimi due o tre decenni, ovvero ai Virno, Negri, Agamben e allo stesso Esposito – trova la sua particolarità in una ben specifica modalità d'utilizzo del pensiero del pensiero di Foucault, in particolar modo del suo concetto di biopolitica. Laddove Foucault e i foucaultiani non-italiani hanno spesso mantenuto la riflessione sui rapporti tra vita e politica al livello di un'indagine empirica di tipo storico e sociologico, i foucaultiani dell'*Italian Theory* hanno dato alla riflessione biopolitica un'inflexione spiccatamente teoretica. Dalla storia delle idee e dei sistemi di pensiero praticata da Foucault, cioè, questi pensatori italiani sono passati all'ontologia, un'ontologia del *bios*, del politico e della loro interazione. La seconda parte della mia tesi sarà allora che, per arrivare a questo trasferimento della categoria di biopolitica dalla storia delle idee alla filosofia teoretica, una parte significativa l'abbia giocata il pensiero di Spinoza, o per meglio dire un certo modo di interpretare e, vuoi anche, di utilizzare Spinoza che dobbiamo al lavoro di Gilles Deleuze e Antonio Negri.

Cominciamo perciò dai tre tratti fondamentali del pensiero italiano che Esposito propone di identificare. Da Machiavelli ad Agamben, il pensiero italiano sarebbe caratterizzato dalla tendenza a quelle che Esposito chiama 1) l'immanentizzazione dell'antagonismo, 2) la storicizzazione del non-storico e 3) la mondanizzazione del soggetto. Presi insieme, questi tre aspetti della filosofia italiana la rendono un pensiero caratterizzato da quella che Esposito

chiama “l’attualità dell’originario” (2010, pp. 23-33). Ciascuna di queste figure di pensiero può essere letta come la rielaborazione di uno specifico concetto della riflessione di Foucault. Il concetto di immanentizzazione dell’antagonismo rielabora la proposta foucaultiana di sostituire una filosofia politica basata sul concetto di sovranità con una che si costituisca in quanto “microfisica del potere” (Foucault, 1977); quello di storicizzazione del non-storico riprende il concetto di genealogia nella specifica formulazione offertane da Foucault nel saggio intitolato *Nietzsche, la genealogia e la storia* (1977, pp. 29-54); il concetto di mondanizzazione del soggetto è una ripresa di quello foucaultiano di morte dell’uomo (Foucault, 1966). Persino il concetto di “attualità dell’originario”, quando si sarà inteso che cosa Esposito intenda con “originario” e “origine”, finirà per apparire non lontano da ciò che Foucault chiamava, ancora una volta sulla scorta di Nietzsche, “ontologia dell’attualità” (Foucault, 1978; 1984a; 1984b). Quando parla di immanentizzazione dell’antagonismo, Esposito fa riferimento ad un tratto del pensiero italiano che trova in Machiavelli, Gramsci e l’operaismo italiano le sue figure principali. Immanentizzazione dell’antagonismo significa che il pensiero italiano immagina l’orizzonte del politico in maniera apertamente alternativa ad Hobbes. Questi – e con lui il filone prevalente della filosofia politica della modernità – vedeva il conflitto e l’antagonismo come appartenenti alla vita dell’uomo nello stato di natura, immaginando che sia intervenuta poi, a mezzo di un patto, una figura sovrana in grado di rimuovere il conflitto e dare origine a un potere legittimo. Il pensiero italiano da Machiavelli a Gramsci ritiene invece che il conflitto sia parte integrante, parte costitutiva dell’ordine, ovvero che non sia né auspicabile né ipotizzabile un ordine che sappia realmente escludere il conflitto. Se ripensiamo al Foucault di *Sorvegliare e punire* (1975) e *La volontà di sapere* (1976), torna alla mente una presa di posizione molto simile. Qualcuno ricorderà il motto foucaultiano secondo cui la riflessione politica è chiamata finalmente a “tagliare la testa al re” (Foucault, 1977, p. 15), ovvero a smettere di ragionare in termini di sovranità e legittimità, altrettanto a smettere di ritenere che il potere s’incarni unicamente o prevalentemente nelle istituzioni politiche dello stato, fluendo in modo per così dire “deduttivo” dal suo centro fino alle periferie. In luogo di questo genere di riflessione, Foucault proponeva di cogliere invece il potere come qualcosa di microfisico, di usare gli strumenti della ricerca storica e sociologica per riuscire a figurarsi l’onnipresenza e multidirezionalità del conflitto. Rapporti di forza e vettori di potere attraversano l’intero corpo sociale, al centro come alle periferie, nei parlamenti così come nelle scuole, nelle carceri e in ogni altra istituzione nella quale ne vada di saperi, poteri e soggetti, e lo fanno in maniere sempre reversibili, sempre multiformi e complesse, dal momento che non si dà mai potere senza contropotere, mai forza senza una forza antagonistica che le risponda. La differenza tra il taglio di Foucault e quello del pensiero italiano identificato da Esposito sta, per così dire, nel punto d’attacco: tagliare la testa al re significava per Foucault iniziare a fare ricerca empirica sulla base di strumenti e concetti nuovi. Per Esposito, così come in modi diversi per Agamben, significa invece approdare a una filosofia politica e ad una ontologia del *bios* che sappiano problematizzare filosoficamente il concetto di sovranità.

Quello di storicizzazione del non-storico è un concetto che per Esposito riguarda in primo luogo l’opera di Vico, ma altrettanto quelle di Vincenzo Cuoco, Leopardi e De Sanctis. Con questa espressione Esposito nomina una riflessione sulla storicità che non è assimilabile né al tradizionale provvidenzialismo cristiano, né alle filosofie della storia che deriverebbero, secondo una nota linea interpretativa, dalla sua secolarizzazione. Di contro all’asse portante della riflessione filosofica sulla storia, il pensiero italiano concepisce a suo avviso la storia come segnata da un rapporto intrinseco con un elemento non-storico, con qualcosa di non completamente assimilabile alla storicità e alla razionalità ch’essa dispiega, ma che nondimeno entra a far parte in maniera latente della storicità stessa e in qualche modo la mina. Esposito parla di questo qualcosa come dell’originario, dell’origine non storica della storia, e lo

caratterizza come la falda vitale, corporea e animale dello storico, contigua talvolta al linguaggio della magia e del mito, talaltra ad uno stato di natura lacerato dal conflitto. Torna alla mente anche in questo caso Michel Foucault, il quale, nel saggio *Nietzsche, la genealogia e la storia*, ricorreva alla lingua tedesca per distinguere nel pensiero di Nietzsche due opposti concetti di origine. La prima figura è l'*Ursprung*, quel tipo di origine che racchiude e rivela la verità essenziale di un concetto, un'idea, un'istituzione; si tratta di un'origine che, una volta conosciuta, rende conto del percorso storico di quell'idea e offre un fondamento legittimo allo stato attuale della sua evoluzione. L'altra, opposta, figura dell'origine è invece quella espressa dalla parola tedesca *Herkunft*, letteralmente estrazione, provenienza, lignaggio. L'*Herkunft* è quel tipo di origine che rivela come ogni cosa sia nata dal suo contrario, da contingenze derisorie e atte a distruggere ogni infatuazione, insomma da un orizzonte di spessa e resistente materialità storica e sociale, il cui riconoscimento mina anziché assicurare, problematizza anziché fondare (1977, pp. 29-54). Ecco, Esposito attribuisce alle filosofie della storia dei filosofi italiani un'attenzione per la *Herkunft* di contro alla *Ursprung*. La differenza è ancora una volta che la riflessione di Foucault era mirata ad una ricerca empirica di tipo storico-genealogico relativa a singole idee e istituzioni (che cosa sono gli ideali ascetici? aveva chiesto Nietzsche; da quando concepiamo la follia come malattia mentale? chiede per esempio Foucault). Esposito, al contrario, sposta l'antitesi tra *Herkunft* e *Ursprung* nel campo della filosofia della storia, applicandola cioè non a singoli fenomeni storici, a singole idee o istituzioni, ma alla storia in quanto tale.

La mondanizzazione del soggetto è il terzo paradigma della filosofia italiana. Il pensiero italiano offre per Esposito un'alternativa alla linea trascendentale che si afferma con Cartesio e Kant. Tale declinazione trascendentale del filosofare è basata su quella che Esposito chiama "logica della presupposizione come struttura costitutiva della soggettività" (2010, p. 30), una logica in virtù della quale il soggetto presuppone se stesso tanto rispetto al mondo dell'esperienza quanto rispetto a se medesimo. Esposito descrive questa figura della soggettività come aporetica, dal momento che "poggia su se stessa, pure essendo, al contempo, la sostanza su cui poggia" (Ibid). Ancora una volta non siamo lontani dall'argomentare foucaultiano, in particolar modo da quelle pagine de *Le parole e le cose* che caratterizzano l'uomo che funge da oggetto delle scienze umane come un "allotropo empirico-trascendentale" (Foucault, 1966, p. 346), come quella strana figura che è al contempo soggetto e oggetto, forma e carne, il fondamento e ciò che viene fondato. Ancora una volta, però, le soluzioni divergono. Se per Foucault mondanizzare il soggetto, e cioè propiziare la morte dell'uomo, significava cimentarsi col compito multiforme delle genealogie dei saperi e dei poteri che ci costruiscono come soggetti, Esposito e la filosofia italiana scelgono ancora una volta una via più spiccatamente teoretica. Scrive Esposito: "Tutto ciò che nel pensiero italiano va nella stessa direzione – dal primato metafisico del corpo in Vico alla valorizzazione del mondo animale in Leopardi – nasce da quella fonte di senso che perfino Gentile ha posto all'origine del 'metodo dell'immanenza'. Immanente [...] è una vita germinante da se medesima, non posposta, o sottoposta, a una figura soggettiva, comunque declinata, anteposta al suo sviluppo" (2010, p. 264).

Trovo che questo spostamento da una ricerca empirica di tipo storico-sociologico all'ontologia sia rilevante tanto per comprendere le sorti del foucaultismo quanto per indagare la natura dell'*Italian Theory*. Tale spostamento teorico, tra l'altro, mi sembra connesso anche con la reazione relativamente tiepida degli esponenti dell'*Italian Theory* dinnanzi all'ultimo Foucault, quello della cosiddetta svolta etica (vedi 2007; 2009). Se infatti leggiamo la svolta etica, il ritorno al pensiero greco-romano e alla filosofia come stile di vita, come un tentativo del Foucault degli anni '80 di risolvere le aporie teoriche accumulate nel decennio precedente – non è detto che questa lettura sia corretta, ma è la lettura che tendo a sottoscrivere – allora ci rendiamo conto dei motivi per cui gli Esposito, Negri e Agamben, i quali cercano questa

soluzione in direzioni estremamente diverse, preferiscono ricollegarsi direttamente al Foucault degli anni '70 e poi tentare di svilupparlo in nuove direzioni.

Questa ontologizzazione della genealogia nietzscheana e foucaultiana passa anche da Spinoza. Il primo passo in questa direzione era stato compiuto già nel 1962 da Gilles Deleuze col suo *Nietzsche e la filosofia*, in cui Deleuze sistematizzava il pensiero del filosofo della volontà di potenza in quella che potremmo definire un'ontologia delle forze basata sulla distinzione tra affermazione e reazione, tra attivo e passivo. Il secondo passo lo dobbiamo pochi anni dopo allo stesso Deleuze, che nel 1968 pubblica *Spinoza e il problema dell'espressione*. Si tratta probabilmente del più occhiuto e accademico tra i testi deleuziani (fu in effetti la sua seconda tesi di dottorato), un testo il cui centro storiografico era la tesi che la filosofia di Spinoza fosse da leggersi in rapporto al pensiero di Duns Scoto. Al netto di un tentativo di rigore e acribia ermeneutica che può sembrare lontano dalle sue abitudini, è chiaro che le riflessioni di Deleuze sul rapporto tra modi e sostanza in Spinoza e sul tema dell'immanenza siano volte ad avvicinare, per quanto possibile, Spinoza alla filosofia della forza ricostruita sei anni prima nel libro su Nietzsche. Quando Deleuze dà così grande importanza all'idea che quella di Spinoza sia appunto un'etica e non una filosofia morale, sta nei fatti cercando di far valere a proposito del suo pensiero quella che non è nient'altro che la distinzione nietzscheana della prima dissertazione della *Genealogia della morale*, la distinzione tra il buono come opposto al Böse, al malvagio, e il buono come opposto allo *Schlecht*, al cattivo.

Il terzo e definitivo passo è costituito dalla pubblicazione nel 1981 de *L'anomalia selvaggia*, il primo della trilogia di testi dedicati da Toni Negri al pensiero di Spinoza, che andrà a includere anche *Spinoza sovversivo* e *Democrazia ed eternità in Spinoza* (Negri, 2006). L'operazione condotta in *L'anomalia selvaggia* – libro che Negri, com'è noto, scrisse letteralmente mentre era in carcere – prosegue la linea deleuziana e al contempo muove dichiaratamente, sin dalla prima pagina della *Prefazione*, in direzione di un'aperta politicizzazione in chiave radicale dell'ontologia spinoziana. Con spavalderia Negri si concede di definire Spinoza un pensatore “ateo e maledetto” e di affermare che per lui “leggere Spinoza ha rappresentato un'esperienza di incredibile freschezza rivoluzionaria” (2006, p. 27). L'idea di Negri è che la metafisica di Spinoza fondi il materialismo moderno, il cui orizzonte è quello di un'immanenza e di un ateismo in cui ogni ordine presupposto all'agire umano e all'auto-espressione dell'essere è negato. Per Deleuze come per Negri, Spinoza è un filosofo dell'immanenza e della pluralità, della liberazione del desiderio come potenza costruttiva, dell'essere come produttività spontanea, incoercibile e priva di gerarchia.

Mi sembra che questo modo di utilizzare Spinoza (e Nietzsche) abbia avuto un ruolo, foss'anche soltanto al livello della mera ricerca di metafore e dispositivi retorici, nel percorso che ha portato *l'Italian Theory* dall'operaismo degli anni Sessanta e dal pensiero di Foucault al lavoro su concetti come quelli di *bios* e *zoe*, di *communitas* e *immunitas*, di singolarità e moltitudine. Senza con questo voler contestare la bontà storiografica del lavoro fatto da Esposito, mi piace notare, in conclusione, quanta varietà e complessità d'influenze teoriche abbia concorso alla formazione di un paradigma interpretativo capace di proporre un nuovo sguardo sull'origine della filosofia italiana.

#### REFERENCES

- Cavell, S. (1981). *Pursuits of Happiness: The Hollywood Comedy of Remarriage*. Cambridge MA: Harvard University Press;
- Cavell, S. (1989). *This New Yet Unapproachable America: Lectures After Emerson After Wittgenstein*. Chicago: University of Chicago Press;
- Cavell, S. (1990). *Conditions Handsome and Unhandsome: The Constitution of Emersonian Perfectionism*. Chicago: Chicago University Press;

- Chiesa, L., & Toscano, A., Eds. (2009). *The Italian Difference: Between Nihilism and Biopolitics*. Melbourne: re.press;
- Claverini, C. (2016). La filosofia italiana come problema. Da Bertrando Spaventa all'*Italian Theory*. *Giornale critico di storia delle idee*, 15-16, 179-188;
- Claverini, C. (2017). The Italian “difference”. Philosophy between old and new tendencies in contemporary Italy. *Phenomenology and Mind*, 2, 256-262. DOI: 10.13128/Phe\_Mi-21124;
- Deleuze, G. (1962). *Nietzsche e la filosofia e altri testi*, Ed. by F. Polidori, Torino: Einaudi 2002;
- Deleuze, G. (1968). *Spinoza e il problema dell'espressione*. Ed. by S. Ansaldi, Macerata: Quodlibet 1999;
- Esposito, R. (2010). *Pensiero vivente. Origine e attualità della filosofia italiana*. Torino: Einaudi;
- Foucault, M. (1966). *Le parole e le cose*. Milano: BUR 2006;
- Foucault, M. (1975). *Sorvegliare e punire. Nascita della prigione*. Torino: Einaudi 2008;
- Foucault, M. (1976). *La volontà di sapere. Storia della sessualità 1*. Milano: Feltrinelli 2008;
- Foucault, M. (1977). *Microfisica del potere: interventi politici*. Torino: Einaudi;
- Foucault, M. (1978). *Illuminismo e critica*. Roma: Donzelli, 1997.
- Foucault, M. (1984a). What is Enlightenment. In P. Rabinow (Ed.), *The Foucault Reader* (pp. 32-50). New York: Pantheon Books;
- Foucault, M. (1984b). Qu'est-ce que les Lumières. *Magazine littéraire*, 207, 35-39;
- Foucault, M. (2007) *L'ermeneutica del soggetto. Corso al Collège de France (1981-1982)*. Ed. by M. Bertani, Milano: Feltrinelli;
- Foucault, M. (2009) *Il governo di sé e degli altri. Corso al Collège de France (1982-1983)*. Ed. by M. Galzigna, Milano: Feltrinelli;
- Gentili, D. (2012). *Italian Theory. Dall'operaismo alla biopolitica*. Bologna: Il Mulino;
- Hardt, M., & Negri, T. (2000), *Empire*. Cambridge MA: Harvard University Press;
- Hardt, M., & Virno, P., Eds. (1996). *Radical Thought in Italy: A Potential Politics*. Minneapolis: University of Minnesota Press;
- Negri, T. (2005). *La differenza italiana*. Milano: Nottetempo;
- Negri, T. (2006). *Spinoza*. Roma: DeriveApprodi.



---

CORRADO CLAVERINI

Università Vita-Salute San Raffaele

c.claverini@gmail.com

---

# DOVE VA LA FILOSOFIA ITALIANA? RIFLESSIONI SULL'ITALIAN THOUGHT

---

## *abstract*

*In recent years, Italian philosophy is having enormous success in the United States, to the point that today we hear more and more of the “New Italian Thought”. The objective of this essay is to analyse this phenomenon from a historical point of view and see how and to what extent interest in Italian thought has grown in the USA, traversing the fundamental stages that have contributed to the acknowledgement of the “New Italian Thought”.*

---

## *keywords*

*italian thought, Roberto Esposito, italian theory, deterritorialization, italian identity*



**1. Dove va la filosofia italiana?**

Attualmente una delle tematiche filosofiche più discusse a livello nazionale e internazionale è quella riguardante l'*Italian Thought*. Con tale espressione inglese (o quella equivalente di *Italian Theory*) – usata diffusamente a partire soprattutto dalla pubblicazione di *Pensiero vivente* (Esposito, 2010) – gli studiosi intendono sottolineare l'interesse sempre crescente per la filosofia italiana in Nord America e, in particolare, negli Stati Uniti. L'opera appena menzionata di Esposito fornisce alcune prove di tale interesse citando tre recenti antologie in lingua inglese: *Recording Metaphysics. The New Italian Philosophy* (Borradori, 1988), *Radical Thought in Italy. A Potential Politics* (Hardt e Virno, 1996) e *The Italian Difference between Nihilism and Biopolitics* (Chiesa e Toscano, 2009). In territorio statunitense vi è sempre stata, nel corso del Novecento, una particolare attenzione per la filosofia europea: prima del grande successo che ha conosciuto l'*Italian Thought*, va segnalata la diffusione della *French Theory* a partire dalla seconda metà degli anni Sessanta e, ancor prima, quella della *German Philosophy*, dovuta all'emigrazione forzata di pensatori come Adorno, Horkheimer e Marcuse durante il nazismo. Dunque, seguendo il resoconto di Esposito, in quella che potremmo definire la circolazione statunitense del pensiero europeo sono individuabili tre diversi "momenti" (rispettivamente tedesco, francese e italiano). Se volessimo limitarci all'analisi del "momento italiano", dovremmo innanzitutto rilevare il fatto che Esposito non è l'unico ad insistere sulla linea interpretativa appena riassunta.<sup>1</sup>

Sono molti a prendere atto del fatto che oggi sia in corso un cambiamento assai rilevante riguardante la percezione che si ha della filosofia italiana all'estero. Se *Pensiero vivente* si apre con la constatazione che, "dopo un lungo periodo di ripiegamento, o quantomeno di stallo, sembra riaprirsi un tempo propizio per la filosofia italiana" (Esposito, 2010, p. 3); Dario Gentili rileva – in maniera simile – il fatto che "è il pensiero di alcuni filosofi italiani a caratterizzare un cambiamento d'egemonia nella filosofia contemporanea" (Gentili, 2012, p. 7). È in atto – continua Gentili – una "ribalta della filosofia italiana nell'attuale dibattito internazionale" (*ibidem*). Infatti, "mentre si lamenta ritualmente l'arretratezza dei nostri studi, i filosofi italiani sfondano in America – non tanto nei dipartimenti di filosofia, ancora dominati dalla linea analitica, ma nell'ambito degli studi politici e sociali, dell'arte e della letteratura, postcoloniali

---

1 Sull'*Italian Thought* cfr., oltre a *Pensiero vivente*, Esposito (2012); Gentili (2012); Contarini e Luglio (2015); Esposito (2015); Gentili e Stimilli (2015); Esposito (2016); Gentili (2016); Maltese e Mariscalco (2016); Gentili (2017); Lisciani-Petrini e Strummiello (2017).

e di genere” (Esposito, 2012).

Di fronte ad una diagnosi così ottimista sarebbe interessante porsi nuovamente una domanda al centro di un libro-inchiesta, pubblicato ormai più di trent’anni fa da Laterza, a cura di Jader Jacobelli: “dove va – se va – la filosofia italiana?” (Jacobelli, 1986, p. VI). In quella sede, vennero date risposte molto varie: “constato che la filosofia sta andando verso una direzione. Ma desidererei che andasse nella direzione opposta” (Norberto Bobbio, in Jacobelli, 1986, p. 28); “va, più o meno, dove è sempre andata (sebbene con passo più stanco e col fiato più grosso). Da oltre un secolo, la filosofia italiana è una provincia del *Reich* filosofico germanico” (Lucio Colletti, in Jacobelli, 1986, p. 39); “come l’impero bizantino nel secolo XV o quello ottomano nel XIX, anche l’impero dei filosofi si è visto spogliare via via delle sue più ricche province” (Luigi Firpo, in Jacobelli, 1986, p. 65); “non solo non sono in grado di dire dove va la filosofia italiana, ma sono persuaso che la filosofia in generale non deve andare in nessun posto: cioè non deve prefiggersi nessuna determinata meta” (Vittorio Mathieu, in Jacobelli, 1986, p. 115); “va, più o meno nietzscheanamente, ‘inerpicandosi su menzogneri ponti di parole, girovagando, trascinandosi attorno su arcobaleni di bugie’” (Antimo Negri, in Jacobelli, 1986, p. 133); “la filosofia italiana oggi si muove con spregiudicata intelligenza all’interno di nuovi modelli di ragione e all’interno di nuove regioni del sapere” (Franco Rella, in Jacobelli, 1986, p. 157); “la filosofia italiana ‘va’” (Carlo Sini, in Jacobelli, 1986, p. 179); “parlerei volentieri di *dipendenza della filosofia italiana da quella straniera*” (Carlo Augusto Viano, in Jacobelli, 1986, p. 200); “il dibattito filosofico italiano è oggi aperto a tutte le correnti del pensiero contemporaneo” (Vincenzo Vitiello, in Jacobelli, 1986, p. 207).

Insomma, trent’anni fa le prospettive degli studiosi interpellati erano molto varie e oggi forse – volendo riproporre il medesimo quesito – sarebbero del tutto differenti. Recentemente, Massimo Ferrari si è posto la domanda sull’andamento della filosofia italiana, rilevando come si possa “almeno constatare che nel nuovo millennio qualcosa si è riequilibrato nello scambio ineguale che a suo tempo ha alimentato lo scetticismo di Bobbio e di Viano” (Ferrari, 2016, p. 296). Infatti, “il complesso di essere una ‘provincia filosofica’ è stato finalmente superato” (ivi, p. 297), anche se – continua Ferrari – la nostra cultura filosofica “non svolge un ruolo trainante nella discussione contemporanea” (*ibidem*). Per i pessimisti, tali affermazioni potrebbero ancora essere motivo di grande sconforto. Per gli ottimisti, al contrario, è possibile che la situazione attuale appaia certamente molto più rosea rispetto a qualche decennio fa. Abbiamo già visto, ad esempio, come gli *Italian Theorists* appartengano alla schiera degli ottimisti. Pertanto, se si chiedesse a costoro “dove va la filosofia italiana?”, una probabile risposta sarebbe che essa va all’estero e, in particolare, negli Stati Uniti. Finalmente la nostra cultura filosofica viene esportata ed è più produttiva che recettiva.

A questo punto, è utile ripercorrere le tappe che hanno fatto sì che oggi si parli diffusamente di *New Italian Thought*. Un quadro sintetico viene fornito nel recente libro di Alessandro Carrera (2017) che giustamente comincia col constatare il fatto che, in America, vi è sempre stata una forte attenzione nei confronti del pensiero italiano. Citando *Dei delitti e delle pene* di Cesare Beccaria, tradotto in inglese nel 1767 con il titolo *On Crimes and Punishments*, Carrera ricorda l’influenza che tale opera ebbe su Thomas Jefferson. Nel suo *Commonplace Book*, il futuro presidente degli Stati Uniti annotò un passaggio in cui Beccaria prese posizione contro le leggi che proibiscono il porto d’armi.<sup>2</sup> A questo si aggiunge il fatto che Jefferson non fu

## 2. Storie di filosofia italiana in Nord America

---

<sup>2</sup> Per il passaggio annotato da Jefferson cfr. Beccaria, 1764, pp. 106-107: “falsa idea di utilità è quella che sacrifica mille vantaggi reali per un inconveniente o immaginario o di poca conseguenza, che toglierebbe agli uomini il fuoco perché incendia e l’acqua perché annega, che non ripara ai mali che col distruggere. Le leggi che proibiscono di portar le armi

influenzato soltanto dall'illuminista lombardo, ma anche da Filippo Mazzei. Amico e vicino di casa del politico americano, Mazzei gli suggerì la frase “all men are created equal” che sarà inserita nella Dichiarazione di Indipendenza del 1776.<sup>3</sup> Occorre infine ricordare come Benjamin Franklin sia stato lettore appassionato della *Scienza della legislazione* di Gaetano Filangieri, con il quale intrattenne anche uno scambio epistolare. Franklin riprese alcune idee del filosofo napoletano riguardanti l'ambito della procedura penale. Tali idee furono poi inserite nel VI emendamento alla Costituzione americana.

Non essendo questa la sede per ricostruire l'intera storia dei rapporti che sono intercorsi fra il pensiero italiano e quello americano, ci limiteremo a ripercorrere brevemente le principali tappe che nel Novecento hanno portato all'odierno dibattito sul *New Italian Thought*. Seguendo ancora il resoconto di Carrera, è necessario a tale scopo ricordare due date simboliche: quella in cui Benedetto Croce fu invitato all'inaugurazione del Rice Institute di Houston (1912) e quella della traduzione in inglese di *Pensiero vivente* di Esposito (2012). In questi cento anni il pensiero italiano ha conosciuto una sempre crescente diffusione. Per quanto riguarda Croce, egli rifiutò l'invito ricevuto dal Rice Institute, ma compose il *Breviario di estetica* affinché fosse tradotto e letto in sua assenza agli studenti americani. Ancora oggi è facilmente reperibile la crociana *Guide to Aesthetics* e la letteratura critica in lingua inglese sul filosofo idealista è abbastanza ampia. Gentile è meno tradotto, anche se non sono assenti studi anglo-americani sul suo pensiero. Tuttavia, è certamente Gramsci il filosofo italiano contemporaneo più diffuso (non solo nei paesi anglofoni, ma in tutto il mondo).

Comunque, se oggi si parla di *Italian Thought*, ciò non è dovuto soltanto alle traduzioni inglesi e allo studio in territorio statunitense dei classici italiani (come, ad esempio, Machiavelli, Bruno, Vico, Leopardi, Croce, Gramsci). Determinante è stato il successo di Umberto Eco negli anni sessanta-settanta e, in particolar modo, quello di Gianni Vattimo negli anni ottanta-novanta. La diffusione del pensiero di Vattimo ha avuto, fra le varie conseguenze, anche quella di una moltiplicazione delle traduzioni inglesi di opere di filosofi italiani come Agamben, Sini, Cacciari, Rella, Cavarero, Perniola e Ferraris. Tuttavia, non sarebbe stato possibile parlare di *Italian Thought* senza il successo mondiale raggiunto da Negri con la pubblicazione – insieme a Michael Hardt – di *Empire* (2000), *Multitude* (2004) e *Commonwealth* (2009). È Negri uno degli autori che – insieme ad Agamben ed Esposito – più di tutti hanno contribuito “a costituire la massa critica indispensabile al riconoscimento dell'*Italian Philosophy* o *New Italian Thought* come soggetto di produzione culturale e oggetto degno di discussione” (Carrera, 2017, p. 24).

Infine, nel 2010, Esposito ha dato il nome di *Italian Thought* a tale fenomeno di crescente diffusione della filosofia italiana all'estero nella sua opera più volte menzionata *Pensiero vivente*. Pubblicato in inglese nel 2012, *Living Thought* è “un libro per il quale ben quattro

---

sono leggi di tal natura; esse non disarmano che i non inclinati né determinati ai delitti, mentre coloro che hanno il coraggio di poter violare le leggi più sacre della umanità e le più importanti del codice, come rispetteranno le minori e le puramente arbitrarie, e delle quali tanto facili ed impuni debbon essere le contravvenzioni, e l'esecuzione esatta delle quali toglie la libertà personale, carissima all'uomo, carissima all'illuminato legislatore, e sottopone gl'innocenti a tutte le vessazioni dovute ai rei? Queste peggiorano la condizione degli assaliti, migliorando quella degli assalitori, non iscemano gli omicidii, ma gli accrescono, perché è maggiore la confidenza nell'assalire i disarmati che gli armati. Queste si chiaman leggi non prevenitrici ma paurose dei delitti, che nascono dalla tumultuosa impressione di alcuni fatti particolari, non dalla ragionata meditazione degl'inconvenienti ed vantaggi di un decreto universale”.

3 Cfr. l'articolo di Mazzei, tradotto da Jefferson e pubblicato un anno prima della Dichiarazione di Indipendenza sulla “Virginia Gazette”: “per ottenere il nostro intento bisogna, miei cari concittadini, ragionar su i diritti naturali dell'uomo e sulle basi di un governo libero. Questa discussione ci dimostrerà chiaramente, che il britanno non è mai stato tale nel suo maggior grado di perfezione, e che il nostro non era altro che una cattiva copia di quello, [...] ma è finalmente venuto il tempo di cambiar costume [...]. Tutti gli uomini sono per natura egualmente liberi e indipendenti. Quest'eguaglianza è necessaria per costituire un governo libero. Bisogna che ognuno sia uguale all'altro nel diritto naturale” (Mazzei, 1775, pp. 496-497).

case editrici si sono contese i diritti” (ivi, p. 43) e che ha provocato molteplici reazioni. Fra critici e sostenitori, il *New Italian Thought* è oggetto di dibattito costante, come dimostrano le numerose pubblicazioni e i convegni dedicati a tale tematica. In tale contesto si colloca la recente fondazione della SIP (“Society for Italian Philosophy”) e la prima conferenza da essa organizzata (“New Italian Thought: Challenges and Responses”) tenutasi a London (Ontario, Canada) presso il King’s University College dal 24 al 26 marzo 2017. Una menzione particolare meritano infine la collana “Sunny Series in Contemporary Italian Philosophy” dell’editore SUNY Press (di tale collana cfr. Benso e Schroeder 2007; Calcagno 2015; Benso 2017) e l’Italian Thought Network (<http://italianthoughtnetwork.com>), una fonte in costante aggiornamento grazie alla quale è possibile rendersi conto empiricamente dei numerosi eventi, pubblicazioni e news riguardanti il pensiero italiano e la sua diffusione nazionale e internazionale. Questo network ha creato una rete capillare di collaborazioni fra gruppi di ricerca operanti in istituzioni accademiche di tutto il mondo (dall’Italia alla Francia, dalla Germania all’Irlanda, dal Canada agli USA).

Dunque, alla domanda sull’andamento della filosofia italiana, gli *Italian Theorists* risponderebbero che essa va all’estero. In particolare, secondo Esposito, la “fuoriuscita da sé” viene presentata come “il tratto più originalmente vivente del pensiero italiano” (Esposito, 2010, p. 16). Tuttavia, occorre subito specificare il fatto che tale caratteristica non è una novità: la “continua deterritorializzazione” (*ibidem*) ha da sempre contraddistinto la tradizione intellettuale italiana. Infatti, sin dalle sue origini umanistico-rinascimentali, quella italiana è stata una filosofia cosmopolita e non nazionale. Inoltre, sempre per mostrare come quello della costante estroflessione sia una peculiarità storica del pensiero italiano, Esposito ricorda la teoria della circolazione di Bertrando Spaventa (1862), secondo la quale la filosofia moderna – nata in Italia con Campanella, Bruno e Vico – si è potuta sviluppare a pieno soltanto all’estero (con Cartesio, Spinoza, Kant e l’idealismo tedesco) a causa della poca libertà degli italiani, oppressi dalla Chiesa cattolica. Secondo quanto viene argomentato in *Pensiero vivente*, il movimento verso il fuori è caratteristica tipica del pensiero italiano, non solo per la sua continua circolazione (come la chiamerebbe Spaventa) all’estero, ma anche per la sua “singolare propensione [...] nei confronti del non filosofico” (Esposito, 2010, p. 12). Infatti, richiamandosi alle tesi espresse da Garin e la sua scuola,<sup>4</sup> Esposito afferma che “sia l’impegno civile che la contaminazione con altri stili di espressione” (*ibidem*) sono una peculiarità storica della tradizione intellettuale italiana. In particolare, “il contenuto del pensiero italiano è ciò che preme al suo esterno, sollecitandolo in qualche modo a uscire da sé per affacciarsi sullo spazio del fuori” (ivi, p. 13). Il mondo della vita storica e politica – quello che appunto viene definito lo “spazio del fuori” – è stato il costante oggetto di interesse da Machiavelli a Vico, fino a Cuoco, Leopardi, De Sanctis e oltre. A sostegno della sua tesi, Esposito cita anche Remo Bodei, il quale ha recentemente affermato che “le filosofie italiane sono pertanto più filosofie della ‘ragione impura’, che tiene conto cioè dei condizionamenti, delle imperfezioni e delle possibilità del mondo, che non della ragion pura rivolta alla conoscenza dell’assoluto, dell’immutabile o del rigidamente normativo” (Bodei, 1998, p. 64).

Secondo Esposito, il carattere non nazionale e cosmopolita, la singolare propensione nei confronti del non filosofico, la continua tendenza alla deterritorializzazione, la costante estroflessione e il movimento verso il fuori sono caratteristiche di lungo periodo del pensiero italiano. Tali peculiarità sono riscontrabili sin dalla prima modernità e, pertanto, occorre chiarire il fatto che, per quanto concerne l’*Italian Thought*, è possibile darne non soltanto una

### 3. Estroflessione e introspezione

---

<sup>4</sup> Cfr., ad esempio, Garin (1947) e Ciliberto (2012).

lettura sincronica, ma anche una diacronica. In particolare, vi è chi per *Italian Thought* intende soltanto la filosofia italiana degli ultimi cinquant'anni e chi, invece, a partire dall'analisi del pensiero contemporaneo, costruisce una genealogia di lungo periodo per mostrare come vi siano alcune caratteristiche costanti e ben riconoscibili che contraddistinguono l'intera tradizione intellettuale italiana dal periodo umanistico-rinascimentale fino ad oggi. Come ormai sappiamo, il secondo approccio è quello di Esposito che, in *Pensiero vivente*, dà una lettura diacronica di *Italian Thought* e che permette di ascrivere la sua opera a quel particolare genere letterario che si interroga su uno specifico ambito dell'identità italiana, ovvero quello che studia l'esistenza di una modalità tipicamente italiana di filosofare. In tale ambito le pubblicazioni non sono molte, dal momento che numerosi sono gli studiosi che mettono in dubbio che sia legittimo parlare di filosofia in termini di appartenenza nazionale. Tuttavia, in Italia più che altrove, tale genere letterario è diffuso, al punto che, a questa tendenza all'*estroflessione* della filosofia italiana, va aggiunta quella all'*introspezione*. In altre parole, è la propensione all'introspezione una delle innegabili costanti storiografiche che accomuna non soltanto le già menzionate tesi di Spaventa e di Garin, ma anche quelle più recenti degli *Italian Theorists* (e, in particolare, di Esposito<sup>5</sup>). In definitiva, è possibile affermare, senza tema di smentita, che la nostra tradizione ha erto a suo motto il "conosci te stesso" socratico. La domanda iniziale – "dove va la filosofia italiana?" – è stata spesso rideclinata in diversi modi (non solo al presente, ma anche al passato e al futuro): dove è andata storicamente? Che cosa dobbiamo aspettarci in futuro? Ma anche: dove si vorrebbe che andasse? In altri termini, il passaggio dal "chi siamo?" al "chi vogliamo essere?" è immediato. Infatti, l'intento di chi si è posto tali quesiti non è mai stato solo descrittivo. Nessuno si è mai limitato all'illustrazione avalutativa dell'andamento della filosofia italiana. In tutti vi è sempre stata la sottintesa volontà di influenzare tale andamento in una direzione ben precisa, consci del fatto che l'identità (filosofica e – più in generale – culturale) italiana – come ogni altra identità – è un prodotto storico che si modifica costantemente nel tempo. Non è inutile ricordarlo contro le sempre risorgenti forme di identitarismo e di nazionalismo che erroneamente concepiscono l'identità e la nazione come qualcosa di naturale piuttosto che di storico e sociale. Sarebbe al tempo stesso opportuno sottolinearlo nella speranza di arginare il processo di progressiva uniformazione e omologazione culturale che contraddistingue l'odierna epoca della globalizzazione. In altre parole, occorre oggi salvaguardare la molteplicità delle culture nazionali, senza però rinunciare al libero dialogo fra di esse. Anche questa semplice e fondamentale verità si può apprendere dalla storia della filosofia e dallo studio della circolazione delle idee nel tempo e nello spazio.

#### BIBLIOGRAFIA

- Beccaria, C. (1764/2014). *Dei delitti e delle pene*. Milano: Feltrinelli;  
Benso, S. (2017). *Viva Voce: Conversations with Italian Philosophers*. Albany: SUNY Press;  
Benso, S. e Schroeder, B. (Eds.) (2007). *Contemporary Italian Philosophy: Crossing the Borders of Ethics, Politics, and Religion*. Albany: SUNY Press;  
Bodei, R. (1998). *Il noi diviso. Ethos e idee dell'Italia repubblicana*. Torino: Einaudi;  
Borradori, G. (Ed.) (1988). *Recording Metaphysics. The New Italian Philosophy*. Evanston: Northwestern University Press;

---

<sup>5</sup> È necessario tuttavia esplicitare il fatto che, nonostante il comune intento introspettivo, le tesi di questi autori sono molto diverse fra loro e per certi versi incommensurabili. Lo stesso Esposito, ad esempio, muove da una critica all'impianto storiografico di Spaventa (e, in particolare, alle categorie tipicamente idealistiche di precorrimiento e inveramento).

- Calcagno, A. (Ed.) (2015). *Contemporary Italian Political Philosophy*. Albany: SUNY Press;
- Carrera, A. (2017). *Benedetto Croce in Texas. Storie di filosofia italiana in Nord America*. Bergamo: Moretti & Vitali;
- Chiesa, L. e Toscano, A. (Eds.) (2009). *The Italian Difference between Nihilism and Biopolitics*. Melbourne: re.press;
- Ciliberto, M. (Ed.) (2012). *Il contributo italiano alla storia del pensiero. Ottava appendice*. Roma: Istituto dell'Enciclopedia Italiana;
- Contarini, S. e Luglio, D. (2015). *L'Italian Theory existe-t-elle?*. Paris: Mimesis;
- Esposito, R. (2010). *Pensiero vivente. Origine e attualità della filosofia italiana*. Torino: Einaudi;
- Esposito, R. (2012, 24 febbraio). Il made in Italy della filosofia. *La Repubblica* (<http://ricerca.repubblica.it/repubblica/archivio/repubblica/2012/02/24/il-made-in-italy-dellafilosofia.html>);
- Esposito, R. (2015). *Problemi del Novecento filosofico italiano*. In O. Grassi e M. Marassi (a cura di), *La filosofia italiana nel Novecento. Interpretazioni, bilanci, prospettive*. Milano-Udine: Mimesis;
- Esposito, R. (2016). *Da fuori. Una filosofia per l'Europa*. Torino: Einaudi;
- Ferrari, M. (2016). *Mezzo secolo di filosofia italiana. Dal secondo dopoguerra al nuovo millennio*. Bologna: Il Mulino;
- Garin, E. (1947/1978). *Storia della filosofia italiana*. Torino: Einaudi;
- Gentili, D. (2012). *Italian Theory. Dall'operaismo alla biopolitica*. Bologna: Il Mulino;
- Gentili, D. (2016). *L'Italian Theory nella crisi della globalizzazione*. In D. Balicco (a cura di), *Made in Italy e cultura. Indagine sull'identità italiana contemporanea*. Palermo: Palumbo editore, pp. 243-247;
- Gentili, D. (2017). *Italian Theory: crisi e conflitto*. In G. Gamba, G. Molinari, M. Settura e M. Coccoresse (a cura di), *Transizioni e cesure di una modernità incompiuta. Tracce di senso in tempo di crisi: studi su Badiou, Florenskij, Hegel, Italian Theory, Laclau, Marx, Nietzsche, Sloterdijk*. Milano-Udine: Mimesis, pp. 71-84;
- Gentili, D. e Stimilli, E. (2015). *Differenze italiane. Politica e filosofia: mappe e sconfinamenti*. Roma: Derive Approdi;
- Hardt, M. e Virno, P. (Eds.) (1996). *Radical Thought in Italy. A Potential Politics*. Minneapolis: University of Minnesota Press;
- Jacobelli, J. (1986). *Dove va la filosofia italiana?*. Roma-Bari: Laterza;
- Lisciani-Petrini, E. e Strummiello, G. (a cura di) (2017). *Effetto Italian Thought*. Macerata: Quodlibet;
- Maltese, P. e Mariscalco, D. (2016). *Vita, politica, rappresentazione. A partire dall'Italian Theory*. Verona: Ombre Corte;
- Mazzei, F. (1775/1970). *Frammenti di scritti pubblicati nelle gazzette al principio della rivoluzione americana da un cittadino di Virginia*. In *Memorie della vita e delle peregrinazioni del fiorentino Filippo Mazzei (1810-1813)*. Milano: Marzorati;
- Spaventa, B. (1862/1908). *La filosofia italiana nelle sue relazioni con la filosofia europea*. Bari: Laterza.









**CeSEP**  
Centro Studi di Etica e Politica



**gender**  
INTERFACULTY CENTRE for GENDER STUDIES

## Topics

*Phenomenology and Social Ontology; Ethics and Political Theory; Cognitive Neurosciences, Philosophy of Mind and Language, Logic; Aesthetics, Metaphysics and History of Ideas.*

## Frequency

*2 issues per year*

## Editor-in Chief

*Roberta De Monticelli (PERSONA)*

## Co-Editors

*Research Centers*

*Roberta Sala (CeSEP)*

*Matteo Motterlini (CRESA)*

*Andrea Tagliapietra (CRISI)*

*Francesca De Vecchi (gender)*

## Faculty

*Claudia Bianchi, Massimo Cacciari, Massimo Donà, Roberto Mordacci, Massimo Reichlin*

## Vice-Editor

*Stefano Cardini*

## Managing Editor

*Francesca Forlè*

## Editorial Team

*Stefano Bacin, Francesca Boccuni, Emanuele Bottazzi, Emanuele Caminada, Francesca De Vecchi, Francesca Forlè, Alfredo Gatto, Giuseppe Girgenti, Barbara Malvestiti, Francesca Pongiglione, Andrea Sereni, Elisabetta Sacchi, Sarah Songhorian, Silvia Tossut, Francesco Valagussa*

## Graphic Design

*Dondina e associati (print version)*

## Graphic Layout

*Direweb (on line version)*

## Web Site Editorial Board

*Stefano Cardini, Francesca Forlè, Sarah Songhorian*

50,00 €