

PHENOMENOLOGY AND MIND

THE ONLINE JOURNAL OF THE RESEARCH CENTRE IN PHENOMENOLOGY AND SCIENCES OF THE PERSON



IUSS Press

Istituto Universitario di Studi Superiori di Pavia



Editor

Roberta De Monticelli

Vice-Editor

Stefano Cardini

Managing Editor

Francesca De Vecchi

Editorial Board

*Francesca Boccuni, Emanuele Caminada, Stefano Cardini, Francesca De Vecchi, Francesca Forlé,
Barbara Malvestiti, Cinzia Prosio, Serena Rizzello, Sarah Songhorian*

Graphic Design

Dondina e associati

Graphic Layout

Likecube snc

PHENOMENOLOGY AND MIND

THE ONLINE JOURNAL OF THE RESEARCH CENTRE IN PHENOMENOLOGY AND SCIENCES OF THE PERSON

NORMS, VALUES, SOCIETY: PHENOMENOLOGICAL AND ONTOLOGICAL APPROACHES

Edited by Emanuele Caminada and Barbara Malvestiti



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

COPYRIGHT

All rights, including translation, reserved. Except for fair copying, no part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means electronic, mechanical, photocopying, recording or otherwise, without the prior written permission of the publishers, or in the case of reprographic reproduction in accordance with the Italian law on copyright. For the online version, the copyright is guaranteed by Creative Commons Licenses.

Research Center in Phenomenology and Sciences of the Person - SCIENTIFIC BOARD

Phenomenology and Cognitive Neurosciences

Stefano Cappa (Università Vita-Salute San Raffaele, Milano)

Michele Di Francesco (Università Vita-Salute San Raffaele, Milano)

Shaun Gallagher (University of Memphis, University of Central Florida, Københavns Universitet, University of Hertfordshire)

Vittorio Gallese (Università degli Studi di Parma)

Andrea Moro (Institute for Advanced Study IUSS, Pavia)

Michael Pauen (Berlin School of Mind and Brain, Humboldt-Universität)

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

Corrado Sinigaglia (Università degli Studi di Milano)

Claudio de' Sperati (Università Vita-Salute San Raffaele, Milano)

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

Phenomenology and Regional Ontologies

Ontology and Language

Lynne Baker (University of Massachusetts, Amherst)

Jocelyn Benoist (Université de Paris 1- Sorbonne)

Elisabetta Sacchi (Università Vita-Salute San Raffaele, Milano)

Ontology and Religious Studies

Paolo Costa (Fondazione Bruno Kessler, Trento)

Giuseppe Di Salvatore (Fondazione Camprostrini, Verona)

Social Ontology

Francesca De Vecchi (Università Vita-Salute San Raffaele, Milano)

Maurizio Ferraris (Università degli Studi di Torino)

Lorenzo Passerini (Università di Milano-Bicocca)

Wojciech Żelaniec (Uniwersytet Gdański, Università degli Studi di Cagliari)

Phenomenology, Norms and Values

Amedeo G. Conte (Università di Pavia)

Mario De Caro (Università degli Studi Roma Tre)

Paolo Di Lucia (Università degli Studi di Milano)

Giuseppe Lorini (Università degli Studi di Cagliari)

Verena Mayer (Ludwig Maximilian Universität München)
Stefano Rodotà (Università degli Studi di Roma “La Sapienza”)
Massimo Reichlin (Università Vita-Salute San Raffaele, Milano)
Roberta Sala (Università Vita-Salute San Raffaele, Milano)

Phenomenology and Theories of Experience

Stefano Besoli (Università di Bologna)
Vincenzo Costa (Università degli Studi del Molise)
Guido Cusinato (Università degli Studi di Verona, Max Scheler Gesellschaft)
Elio Franzini (Università degli Studi di Milano)
Dieter Löhmar (Universität zu Köln)
Paolo Spinicci (Università degli Studi di Milano)

Phenomenology and Education

Massimiliano Tarozzi (Università degli Studi di Trento)
Vanna Iori (Università Cattolica Sacro Cuore, Piacenza)
Daniele Bruzzone (Università Cattolica Sacro Cuore, Piacenza)

INTERNATIONAL REVIEWERS OF THIS ISSUE

Carola Barbero (Università degli Studi di Torino)
Clotilde Calabi (Università degli Studi di Milano)
Amedeo G. Conte (Università degli Studi di Pavia)
Paolo Costa (Fondazione Bruno Kessler, Trento)
Francesca De Vecchi (Università Vita-Salute San Raffaele, Milano)
Paolo Di Lucia (Università degli Studi di Milano)
Roberto Mordacci (Università Vita-Salute San Raffaele, Milano)
Corrado Roversi (Università degli Studi di Bologna)
Thomas Szanto (University College Dublin)
Francesco Valerio Tommasi (Fondazione Bruno Kessler, Trento)
Ina Marie Weber (Husserl Archiv – Universität zu Köln)

CONTENTS

INTRODUCTION

<i>Emanuele Caminada (Universität zu Köln), Barbara Malvestiti (Università degli Studi di Milano)</i> Norms, Values, Society: A Brief Phenomenological Overview	10
------------------------------------------------------------------------------------------------------------------------------------------------------------------------	----

SESSION 1. NON-INSTITUTIONAL LIFE

<i>Andrea Zhok (Università degli Studi di Milano)</i> A Phenomenological Discussion of Early Imitation	26
-----------------------------------------------------------------------------------------------------------	----

<i>Marco Spina (Université Paris-Sorbonne)</i> Norm and Normality, Starting from Merleau-Ponty	36
---------------------------------------------------------------------------------------------------	----

<i>Joseph Kisolo-Ssonko (University of Sheffield)</i> Love, Plural Subjects & Normative Constraint	46
-------------------------------------------------------------------------------------------------------	----

<i>Neil W. Williams (University of Sheffield)</i> Against Atomic Individualism in Plural Subject Theory	56
------------------------------------------------------------------------------------------------------------	----

<i>Alessandro Salice (Universität Wien)</i> Social Ontology and Immanent Realism	68
-------------------------------------------------------------------------------------	----

CONTENTS

SESSION 2. INSTITUTIONAL LIFE

<i>Lorenzo Passerini Glazel (Università degli Studi di Milano-Bicocca)</i> Institutional Ontology as an Ontology of Types	78
<i>Wojciech Żelaniec (University of Gdańsk)</i> On the Constitutive Force of Regulative Rules	92
<i>Guglielmo Feis (Università degli Studi di Milano), Umberto M. Sconfienza (Tilburg University)</i> Challenging the Constitutive Rules Inviolability Dogma	102
<i>Emanuele Bottazzi, Roberta Ferrario (Istituto di Scienze e Tecnologie della Cognizione di Trento)</i> Appearance Counting as Reality? Some Considerations on Stability and Unpredictability in Social Institutions	112
<i>Gaetano Albergo (Università degli Studi di Catania)</i> Does Ontogenesis of Social Ontology Start with Pretence?	120

CONTENTS

SESSION 3. ETHICAL-POLITICAL LIFE

<i>Massimo Reichlin (Università Vita-Salute San Raffaele, Milano)</i> The Neosentimentalist Argument Against Moral Rationalism: Some Critical Observations	130
<i>Christian Blum (Universität zu Köln)</i> Determining the Common Good: A (Re-)Constructive Critique of the Proceduralist Paradigm	140
<i>Roberta Sala (Università Vita-Salute San Raffaele, Milano)</i> Reasonable Values and the Value of Reasonableness. Reflections on John Rawls' Political Liberalism	150

SESSION 4. SPECIAL ISSUE

<i>Emanuele Caminada (Universität zu Köln), Barbara Malvestiti (Università degli Studi di Milano)</i> An Interview with Martha Craven Nussbaum: Political Emotions. Why Love Matters for Justice	160
<i>Valentina Bambini, Cristiano Chesì, Andrea Moro (Istituto Universitario di Studi Superiori di Pavia)</i> A Conversation with Noam Chomsky: New Insights on Old Foundations	166

INTRODUCTION

Emanuele Caminada, Barbara Malvestiti
Norms, Values, Society: A Brief Phenomenological Overview

EMANUELE CAMINADA
Universität zu Köln
emanuele.caminada@gmail.com

BARBARA MALVESTITI
Università degli Studi di Milano
barbaramalvestiti@virgilio.it

NORMS, VALUES, SOCIETY: A BRIEF PHENOMENOLOGICAL OVERVIEW

abstract

In our paper we introduce three main areas of social ontology that correspond with the sections structuring the current issue of “Phenomenology and Mind”: non-institutional life, institutional life and ethical-political life. We argue three points about these areas, which are represented in the accounts published in this issue: levels of social life and reality; normative levels of life and reality; hierarchical levels of life and reality. Finally, we introduce two interviews, to which the last special section of the issue is devoted. The interviews were conducted by the editors with Martha Craven Nussbaum and by Valentina Bambini, Cristiano Chesi and Andrea Moro with Noam Chomsky.

keywords

Non-institutional life, institutional life, ethical-political life, norms, values

- 1. Our Purposes** The current issue (vol. 3) of “Phenomenology and Mind” deals with social facts. Social facts include many different subjects: persons, groups, norms, values, political systems, economical powers, etc., that are the domain of different theoretical and practical disciplines, such as social and developmental psychologies, empirical sociologies (both quantitative and qualitative), political sciences, legal theories, ethics, economics, etc.

What exactly are phenomenological and ontological approaches to these subjects?

Social facts have historically been the subject of several phenomenological studies: in the early German movement (Max Scheler, Adolph Reinach, Edith Stein, Alexander von Hildebrand, Nicolai Hartmann, Gerda Walther, Theodor Litt, Herbert Spiegelberg, Felix Kaufmann); and its development in Europe (in Spain José Ortega y Gasset; in France Maurice Merleau-Ponty, Jean-Paul Sartre, Emmanuel Levinas, Paul Ricoeur; in Italy Norberto Bobbio, Enzo Paci; in Germany Bernard Waldenfels; in Brasil Antônio Luís Machado Neto); and in the United States (Alfred Schutz, Peter Ludwig Berger, Thomas Luckmann). Phenomenological sociology is now recognized as a school of its own in contemporary social sciences.

On the other hand, ontological approaches to social facts declined in popularity not only in postmodernist culture, but also within phenomenological sociology itself, since one of its most famous pioneers, Schutz, disregarded the ontological research of the early phenomenological movement because of their static essentialism: i.e., the belief in universal and constant structures of social entities independent of the dynamics of the concrete subjectivities of the life-world. Contemporary social and cultural studies presuppose the anti-essentialist refusal of constant features, for example, in the domain of group structures, sex orientation and gender, individual and collective identities, normality and pathologies¹.

¹ Some phenomenologists (like Scheler) spoke about individual essences, such as the essence of a nation or a person, going far beyond the ontological assessment that personhood implies absolute individuality (*haecceitas*). His pretension to have direct access to the metaphysical obscure individualities of Germany, Britain, or Europe cast a cloud over his bright ontology of the individual person, justifying more skeptical approaches in cultural studies and Schutz's mistrust of ontology. Husserl's methodological distinction (finally edited in Hua XL) between ontology and the essence of individuals, however, stressing that eidetic, the ontological study of essences (greek '*eidōs*', '*eide*'), deals with classes and not with individuals, as well as his distinction between ontological eidetics and monadological metaphysics could reset long-lasting prejudices about the role of ontologies in cultural and social sciences.

Still, in the last few decades, apart from these influential debates, analytical philosophers began to struggle with the (formal) nature of social facts, looking for their constitutive features, their invariants, and their properties and trying to systematize the results of these meticulous analyses into a particular branch of general metaphysics: social ontology. In doing so they found unexpected echoes among scholars working on and furthering the phenomenological projects, through the pioneering work of the “Seminar for Austro-German-Philosophy” and the reassessment of “realistic phenomenology” (especially Kevin Mulligan and Barry Smith).

In the last three years, the Research Center in Phenomenology and Sciences of the Person has invited several scholars working in this field to discuss their current research. The discussions were held in the lecture halls of the San Raffaele University (Milan), and on the virtual platform of this research laboratory (www.phenomenologylab.eu). The main goal of the lab is to give voice to an authentic phenomenological spirit in both its analytic (in the sense of conceptual clarity and of the attention for formal logic and ontology) and synthetic (in the sense of openness to the best of the philosophical traditions, including the contemporary intellectual debates and the material or regional ontologies). The best contributions on social ontology that preceded and followed the Spring School 2011, and the International Conference Making the Social World, devoted to John R. Searle’s *Making the Social World. The Structure of Human Civilization* (2010), are collected in this issue of the journal under the title “Norms, Values, Society: Phenomenological and Ontological Approaches”.

We speak about phenomenological *and* ontological approaches both in the disjunctive and conjunctive sense, because this issue testifies to ontological approaches within phenomenological and other different philosophical frameworks (e.g. Plural Subject Theory, Speech Act Theory, Constitutive Rules Theory, Theory of Justice) and because, the contemporary dialogue between phenomenology and social ontology occurs mainly on genuine ontological levels, although phenomenological approaches to social facts do not only consist in phenomenological ontology.

Phenomenological approaches to social facts can in fact be divided into three main areas (Nam-Im-Lee 2006): first, empirical phenomenological sociology, concerned with the qualitative description of social facts in the first and second person perspectives (e.g. Berger’s *Sociology of Religion* or Harold Garfinkel’s *ethnomethodology*); second, ontological phenomenological sociology as a kind of regional ontology dealing with the essential structures of social realities (Reinach, Stein, Walther, Hartmann, Spiegelberg, Kaufmann, Bobbio); and third, “transcendental” or constitutive phenomenological sociology, which finally aims to clarify the condition of the possibility of social reality and its structures².

The main “axiom” of all these branches of phenomenological sociology is that social facts are products of intentionalities, said by Husserl to be constituted in lived experience, and by Searle to be mind-dependent. This general assumption is nowadays widely shared in the contemporary ontological debate (although not in wider cultural and social studies), with particular attention to the constitutive role of collective intentionality and its different modes.

The working hypothesis of our Lab was that the individual person is the last bearer of properties in the ontological region of the social. She is individuated in her intentional positions and attitudes (both

² Nam In-Lee (2006) claims, that we don’t have any phenomenological attempts of transcendental sociology. Maybe Schutz and his heirs Berger and Luckmann (Berger and Luckmann 1966, Schutz and Luckmann 1975/1984) could be considered as transcendental sociologists, although they didn’t. Yet since they looked for invariant but dynamic structures of the life-world, and Husserl considered the description of these the phenomenological way to transcendental philosophy, as the research on the condition for the possibility for something (in our case, social facts), we can consider them in these frame. Thereby Berger and Luckmann’s metaphysical sympathies for constructionism could be revisited within Husserl’s approach to the realism and constructivism debate.

theoretical, axiological and practical) toward her social environment, which is necessary, but not a sufficient condition for her concrete personal development. On one hand, she relies on her natural, biological and psychological faculties, which, together with her material and social environment, give her the possibility to develop her individual personality in early childhood socialization. On the other hand, the mature and autonomous flourishing of her being-person requires the free capability to emerge from the level of her social environment through the execution of spontaneous and rational acts, both shaping her individuality as well as offering her personal contribution to social and institutional life. Personhood, therefore, is at the same time rooted in and transcending sociality. Moreover, the individual person, as the last bearer in the region of social ontology, does not coincide with her natural and psychological support, although she is founded on it. As highlighted by Norberto Bobbio (1948), a human being is definable by three traits, which coincide with its progressive levels of individuation: human being as natural being or biological *individual*, human being as social being or *socius*, human being as personal being or *person*. Curiously, positive law has historically recognized (although not everywhere) the correlative rights in the opposite order: firstly, civil and political rights; then social and economic rights; and finally, contemporary ecological rights as well as biological or psychological or even cultural self-determination rights, as Jeanne Hersch stressed (1990).

Further, individual persons are not the only objects of the ontological region of the social, which is inhabited also by organizations consisting of a plurality of persons (such as clubs, states, churches, universities...), immaterial objects (such as promises, marriages, juridical persons, titles, ...), and social subjects (such as documents, money, monuments and meeting halls for clubs, parliaments and governments, churches and universities), whose material founders cannot explain their irreducible social meaning, although the constitution of social objects is limited by materiality (Roversi 2012, Terravecchia 2012). Hartmann (1933) defined these three domains as the domain of the person or subjective mind (*Person bzw. subjektiver Geist*); the domain of the common mind or objective mind (*Gemeingeist bzw. objektiver Geist*); and the domain of the cultural objects or objectified mind (*Kulturwerke bzw. objektivierter Geist*).

2. Three Topics Three main topics emerge from the accounts presented here. They correspond to three main areas of social reality:

1. Non-institutional life
2. Institutional life
3. Ethical-political life

Therefore, the issue is structured in three sections corresponding to these areas. We will argue for three points about the aforementioned areas, which are supposed to be shared by the accounts presented in this issue:

- (i) They are levels of social life and reality
- (ii) They are normative levels of life and reality
- (iii) They are hierarchical levels of life and reality

Let us present the main features of each point.

2.1 Social Life vs. Natural Life The contemporary ontological debate presupposes a sharp distinction between natural and social facts. Phenomenological ontologies recognized in nature and “social mind” (*Gemeingeist*) two cardinal domain of reality. The general ontological and metaphysical question about the relation between

these two domains transcends the special ontological topic of the present issue. Some scholars, like Searle, embrace naturalism for both domains, others see in naturalism an approach that is consistent only with the domain of nature. Phenomenology is more sympathetic to the relativization of the universal claim of naturalism. Husserl explored nature and mind as different forms of reality given to corresponding attitudes, the naturalistic and the personalistic (Hua IV). Finally, he viewed both attitudes as rooted in the concrete (and social) life-world. Thus, his ontology sees a direct foundation between natural and social life, while the objects of natural sciences (such as atoms, waves, energies, neurons, dna-informations, etc.) have to be clarified in the scientific process that led from every-day experience to the conceivability of these objects (Hua VI).

Without taking a position on these different approaches, we stress the following shared conviction of both transcendental phenomenology, as well as phenomenological and contemporary ontologies: that non-institutional life, institutional life, and ethical-political life are levels of *social* life. Social life is characterized by two cardinal features, which distinguishes it from both natural life, and from non-social intentional life:

1. In contrast to natural life, social life needs *the intentionality of individuals* in order to exist. It is mind-dependent.
2. In contrast to non-social intentional life, social life needs, in order to exist, the *intentionality of a plurality at least of two individuals*³. It is plural-minds-dependent.

The question of whether social ontology should admit intentional but not social forms of life is to be left open⁴.

Non-institutional life, institutional life, ethical-political life are not only different areas of social life, they are also *normative* levels of life. What we mean is that *social reality as such is a normative dimension of life*.

2.2 Social Life as a Normative Level of Life

In particular, we make two claims:

1. Social reality has its own eidetic laws, which differ from the causal laws of nature and the motivational structures of mere individual minds (if they exist).
2. Social reality is characterized by different types of normativity, which vary depending on the level of social life we meet⁵.

Non-institutional life, institutional life, and ethical-political life are not only levels of *social* life, nor are they only *normative* levels of life. They are also *hierarchical* levels of life: non-institutional life, institutional life, and ethical-political life are levels of social life in ascending order of complexity. This implies that social-ontology has also to answer the formal-ontological question about the type of relation among its components. In the contemporary debate, different solutions have been proposed to formalize the inner hierarchy of social reality (e.g. Supervenience Theory, Constitution Theory, Emergentism). We distinguish at least three hierarchical domains of social life: non-institutional, institutional life and ethical-political life.

2.3 Hierarchical Levels of Social Life

3 As Francesca De Vecchi says, social entities do not depend on solitary intentionality; they involve “heterotropic intentionality” (De Vecchi 2012, pp. 17-18).

4 Gilbert claims for such a position, Williams addresses in the present issue his criticism to her opinion.

5 See Zaibert&Smith (2007).

Non-institutional life is the basic level of social life. It is the domain of social life precedent or besides its institutional normation. The first section of this issue is devoted to the topic of non-institutional life.

Institutional life is the level of social life that is characterized by the phenomenon of norm *stricto sensu*: it is the social life in the typical institutional forms. The second section of this issue is devoted to the topic of institutional life.

Ethical-political life is the level of social life that implies the translation of values and duties in political and juridical systems. It represents, in some sense, the apex of the phenomenon of normativity. It is concerned with the level of both social and institutional life regarded in their ontological, deontic and axiological components. Meta-ethical and political topics are discussed in the third section.

3.1
First Topic:
Non-Institutional
Life

The first section presents contributions exploring the social world besides its institutional types, focusing on early imitation (Zhok), the embodied constitution of normality (Spina), the relations between individual persons and personal groups (Ssonko, Williams), and the nature of cultural objects (Salice).

Zhok attempts in *A Phenomenological Discussion on Early Imitation* to apply Husserl's phenomenological approach to early imitation, makes an interesting comparison with empirical researches in developmental psychology (especially by Meltzoff), and questions their theoretical premises concerning the nature of the mind. Finally, he sketches a phenomenological theory about embodied access to feelings and expressions, stressing the relevance of rhythmic structures of experience for the attunement of interpersonal fields that enable the first steps of socialization.

Spina's paper *Norm and Normativity. Starting from Merleau-Ponty* shows the conceptual tension and ambiguity of Merleau-Ponty's description of normality and norm by analysing his phenomenology of perception and facing the task of abnormality.

Kisolo-Ssonko and Williams both work against the background of Gilbert's Plural Subjects Theory⁶. Kisolo-Ssonko deals with the social-ontological relevance of love. Following Westlund's account, he claims in *Love, Plural Subjects and Normative Constraints* that lovers become a plural subject. But if Westlund refuses Gilbert's theory of a direct normative constraint claiming that love liaisons are much more flexible, Kisolo-Ssonko argues against this rejection, still claiming for a revision in Gilbert's account, and suggests the distinction of various levels of possible identification of the lovers with the plural subject they form.

In *Against Individualism in Plural Subject Theory* Williams tests the plausibility of Gilbert's assumption that non-social individuals can autonomously form a plural subject by applying Sperber and Wilson's Relevance Theory. He shows that communication already presupposes some forms of communality among the players. Therefore, he concludes that Plural Subject Theory needs the notion of an already socialized individual as its primary methodological unity, rather than the atomic individual. This conclusion implies that Gilbert's Plural Subject Theory can be maintained as a sub-regional description of social-ontological group phenomena, but that we need a more comprehensive theory in order to describe socialization itself, since only socialized persons can join in the form of plural subjects.

Salice draws in *Social Ontology and Immanent Realism* the distinction between social and cultural facts, claiming that, while the former are culturally universal, the latter are individuated in

⁶ Also in phenomenology we find descriptions of superindividual subjects or plural subjects: Scheler's term *Gesamtperson* (Scheler 1913; 1916), Stein's *überindividuelle Persönlichkeit* (1922), and Husserl's *Personalität höher Ordnung* (Hua XIV).

singular cultures. Salice suggests that cultural objects, being context-dependent, have to be seen as being “immanent objects” of collective beliefs of the members of the contextual culture. Thus, his contribution, after a historical confrontation with Brentano’s position, gives insight into his current work on a theory of cultural entities in the framework of what he calls a “pseudo-Brentanian immanent realism”. Paying particular attention to Searle’s theory of social construction, he tries to show its limits regarding this particular form of cultural objects, and to sketch a way to explain these phenomena.

The second topic, “Institutional Life”, is concerned with the level of social life characterized by the phenomenon of norms or rules *stricto sensu*.

Although the distinction between the institutional level of social life and the non-institutional level of social life is not always adequately pointed out in the literature, we consider it very important.

According to us, the distinction between the institutional level of social life and the non-institutional level of social life is illuminated by the fundamental distinction between *nomic regularity* and *anomic regularity*, pointed out by Amedeo G. Conte (2004, 2011)⁷. Anomic regularity is the regularity of actions which are performed regularly, but without a rule – anomic regularity is a regularity without rules – nomic regularity is the regularity of the actions which are performed regularly with a rule – nomic regularity is a *rule*-related regularity.

While the non-institutional level of social life is characterized by anomic regularity, the institutional level of social life is characterized by nomic regularity. Typical examples of anomic regularity are iterative imitation acts such as what happens in early imitation; cooperative actions characterized by collective intentionality but not by status functions assignment (e.g. in tribal hunting or in a tribe’s use of a boundary wall, Searle 1995); uses and customs. These last ones have social normative powers, but they are not explicitly governed by rules: some families use to have a particular dinner on a particular day, partners use to devote themselves particular moments of the day. Uses and customs have their own anomic normativity, since they involve implicit commitments, which can be expressed in the form “we are used to A” or “A is normal”⁸. Therefore it is important to distinguish in the domain of normativity between normality (anomic) and norms (nomic). If not normal behaviors happen, it could be that a rule is imposed: regularities become rules. For example if kids do not help to clean after dinner (given the case they did it normally in ideal “innocent” times), parents could institutionalize the use (maybe through sanctions or *premia*).

On the opposite side, typical examples of nomic regularities are the habit of a community to pray *for the sake of a rule*; the habit of a person to remove the cap entering the church because of the fear of social blame or the habit of a person to observe traffic rules because of the fear of police sanctions (to perform action *according to a rule*)⁸; the habit of the cheat to cheat as *a function of a rule* – a rule of a game – to which he does not conform his behavior, *etc.*⁹.

⁷ Conte highlights three forms of nomic regularity – *nomonomic regularity*; *nomophoric regularity*; *nomotropic regularity* – and he distinguishes them from *anomic* or *nomological* regularity. *Nomonomic regularity* is the regularity of actions which are performed for the *sake of a rule*. *Nomophoric regularity* is the regularity of the actions which are performed *according to a rule*. *Nomotropic regularity* is the regularity of the actions which are performed as *a function of a rule*.

⁸ Distinction between *nomonomic regularity* and *nomophoric regularity* is instantiated by Kantian distinction between morality and legality: while *nomonomic regularity* is the regularity of actions which are performed for the “goodness” of the rule itself (*for the sake of a rule*), *nomophoric regularity* is the regularity of the actions which are performed not necessarily for the goodness of the rule (*according to a rule*). As highlighted by Amedeo G. Conte, while “for the sake of a rule” translates the Kantian “*um des Gesetzes Willen*”, reformulated by Conte as “*um einer Regel Willen*”, “*according to a rule*” translates the Kantian “*gemäss dem Gesetze*”, reformulated by Conte as “*gemäss einer Regel*”.

⁹ While non-institutional level of social life is a case of anomic regularity – it would be cases of anomic regularity that are not examples of social life, such as the habit of a person to wash the teeth before sleeping – institutional level of social life exhausts the cases that are characterized by nomic regularity.

3.2 Second Topic: Institutional Life

The fact that non-institutional level of social life is characterized by anomic regularity does not therefore imply the absence of normativity. As much as they share normativity, both non-institutional levels of social life and institutional levels of life have *deontic powers*: they both imply duties, claims, and commitments. But while non-institutional duties are motivated by values and commitments, institutional ones are primarily ruled by norms.

According to Searle's seminal insights, institutional life is characterized by constitutive rules and status functions, while non-institutional social life is not. Let us clarify this distinction. The institutional level of social life contemplates both constitutive rules and status function. Constitutive rules are those rules which create the possibility of the entity they rules (Searle 1969). For example, the rules of chess or rules which regulate Parliamentary acts are those rules without which the game of chess or Parliamentary acts would not exist. They differ from regulative rules, which simply rule behaviors pre-existing to them, such as the rule that prohibits smoking. According to Searle, social entities are essentially status functions, i.e. entities created by constitutive rules or, according to some revisions of his theory, by Status Function Declarations (Searle 2010).

Now, the existence of *institutional* facts, such as "the President of United States", the "owner of a property in Berkeley", "the 20 Euro note" depends on constitutive rules and Status Function Declarations. On the contrary, the existence of social facts, such as love, friendship, solidarity phenomena and so on, seems not to depend on constitutive rules and Status Function Declarations. It sounds very strange to argue for the existence of rules in absence of which love, friendship, solidarity phenomena and so on, do not exist. These phenomena seem to be more spontaneous than institutional ones. Similarly, it seems difficult to consider such phenomena as Status Functions: lovers or friends are supposed to be for each other something different from mere Status Function.

The distinction between *constitutive rules* and *regulative rules* gave rise to a great deal of research on the topics of constitutive rules, such as, in Italy, those of Gaetano Carcaterra and those of Amedeo G. Conte's School¹⁰, which distinguished between *eidetic constitutive rules* and *anankastic constitutive rules* (Conte 2007). While eidetic-constitutive rules *are* a necessary condition for their object, anankastic-constitutive rules *impose* a necessary condition for their object. Eidetic-constitutive rules *create* the *type* of their object, such as the type of the game "chess" and its praxems (example of rule: "one cannot castle when the king is under the check"); anankastic-constitutive rules *determine* the *tokens* of pre-existing types, such as the tokens of the type "wills" (example of rule: "wills ought to be signed by the Testator" (Conte 2001, 73).

While eidetic-constitutive rules are context-independent, anankastic-constitutive rules are context-dependent. Eidetic-constitutive are typical of games, anankastic-constitutive rules are typical of juridical systems.

Nevertheless, there is a level of human juridical systems at which context-independent rules are expected to be recognized. It is the level of common values. At this level neither constitutive rules nor regulative rules are enough to define human juridical system: an inquiry into values is open.

The second section of this issue, "Institutional Life", regards the question of *what essentially characterizes institutional entities*. Differently from the previous section, which primarily involves research on

¹⁰ The distinction between *regulative rules* and *constitutive rules* (Searle 1969) have an important precursor in the Polish philosopher Czesław Znamierowski' s distinction between "coercive norms" (*normy koercyjne*) and "constructive norms" (*normy konstrukcyjne*). See Znamierowski (1924).

modalities of the constitution of non-institutional life, the present section does not correspondingly involve research on modalities of the constitution of institutional life. Contributions of this section welcome institutional reality as already constituted and focus on *what essentially characterizes it*.

We would like to mention each contribution, which points out very relevant aspects about institutional reality, as well as about the phenomenon of constitutive rules. Some of them are also critical of the classical paradigm of constitutive rules and of Searle's institutional reality account.

Lorenzo Passerini's *Institutional Ontology as an Ontology of Type* formulates the hypothesis that the ontology of institutional phenomena is primarily an *ontology of types*, while the ontology of natural phenomena is primarily an *ontology of tokens*. In detail, the author shows four essential characteristics of institutional phenomena.

Wojciech Żelaniec's *On the Constitutive Force of Regulative Rules* argues for the constitutive force of *regulative* rules, although he maintains that they remain a *genus* of their own. The author points out four main reasons in virtue of which regulative rules have a constitutive import; one of these is that regulative rules define new forms of behavior, the behavior compliant with them. The behavior is not new in itself, but it is new depending on what the motives are: abstaining from smoking since we have no wish to smoke is not "doing the same thing" whether we act out of respect for the law.

Guglielmo Feis and Umberto Sconfienza's paper *Challenging the Constitutive Rules Inviolability Dogma* tries to challenge the dogma of the inviolability of constitutive rules. The authors develop an interesting *parametrical* approach to *constitutive rules*, according to which it is possible to violate a constitutive rule. Finally, they also introduce two different ways of exiting a game.

Emanuele Bottazzi and Roberta Ferrario's paper *Appearance Counting as Reality? Some Considerations on Stability and Unpredictability in Social Institutions* argues for a revision of Searle's concept of objectivity. According to their view objectivity has a crucial role in institutional systems as a requirement, but objectivity is not a necessary condition for them to be institutional systems.

Finally, Gaetano Albergo's *Does Ontogenesis of Social Ontology Start with Pretence?* considers the case of pretend play. The author claims that, like the other games, pretend play depends on status function assignment and constitutive rules. Nevertheless, in order to consider the new scenario as a possible world and to abandon the *natural* necessity implicit both in the stipulation and in successive possible implications, we should admit that rules of pretend-inference have a robustly objective status.

The third topic, "ethical-political life", concerns the level of translation of values and rules in political and juridical systems. It represents the apex of the phenomenon of normativity. In what sense?

Following Husserl's masterful universal-ontological project values and norms are integral parts of both formal and material ontologies. Formal logic and ontology are therefore to be integrated by formal ethic and formal deontics (see Hua XXX, Mulligan 2004). Material ontologies should be further integrated by corresponding material ethics (Scheler, Hartmann) and by the desideratum of scientific policies or (*Praktik*, see Hua XXX). Regarding the region of social ontology research a lot is still to be done on the systematization of its peculiar substrates, relations, wholes and parts and on their respective axiological and deontic characters.

Values are not a specific object of social ontology. Other regional ontologies show values: unorganic

3.3 Third topic: Ethical-Political Life

material and natural life, for example, can carry own values. Every realm of reality can be a bearer of values. Concerning the nature of values, phenomenology stands within metaethics for a realistic, objectivistic and cognitivistic position. There are plausible phenomenological reasons to argue for an enlargement of the rationality sphere to emotional life, which provides the possibility of an ethics of values, integrating the Kantian formal rational view of ethics (see Scheler 1913-1916, Hartmann 1925). Although values, as such, are not specific to social ontology, to this region correspond specific regional values, such as sincerity, fidelity, fairness, human dignity, etc. and their corresponding normative features, such as duties, commitments, claims. Furthermore, specific values emerge at the level of institutional life: institutional life coincides with the social life step in which the constitution of norms or rules – obligations, duties, rights, permissions, prohibitions, authorizations, licenses, etc., is given. Norms are, in opposition to values, specific social ontological objects. The constitution of norms could be given without any relation to an order of values, like the rules of a game: norms are ideally independent from values, as Hume teaches. Otherwise, peculiar duties correspond to every value, since the nature of values founds affordances in the form of “to-be-ought”, and therefore motivates and justifies actions. If norms are independent from values, duties are not.

Nevertheless, there is a level of social reality in which the inquiry into the truth of norms, their relation to duties stands out: it is the level of life together, i.e. the ethical-political level, which demands the translation of values and claims into norms. It is the level of life in which the fundamental principles of our Constitutions, the political debates about the reasons for both regulative and constitutive rules are involved. If we are willing to consider values not as what enjoys critical immunity, but as what is liable to discussion and review, we can rely on a possible inquiry into the truth of norms¹¹. In this sense, another big task is open for social ontology: not only the study of relation between values and duties, but also the study of relation between values and norms. Two main research areas for social ontology stand out: “axiology”, which contemplates the study of values and their grounds – “what make good things good” – (Spiegelberg 1947); and “praktology”, which contemplates not only duties and claims, but also rights (Spiegelberg 1933). In this view, rights are conceived as the institutionalized recognition of claims, which are then motivated by values (Spiegelberg 1939).

The contributions of this section have the merit to reconstruct the classical paradigm of rationalism, in ethics and politics, and to criticize or integrate it, providing interesting directions of research, worth of developments.

Massimo Reichlin’s *The Neosentimentalist Argument against Moral Rationalism: Some Critical Observations* deals with the clarification of neosentimentalist approaches to metaethics, according to which our moral judgments are the expression of our sentiments and affective reactions, without any intentional and cognitive character. After having illustrated a sophisticated formulation of the neosentimentalism, which make it possible to overcome weaknesses of a naive formulation of neosentimentalism, the author still suggests that such a formulation weakens neosentimentalism attacks against moral rationalism. Then the author outlines a personal proposal which tries to integrate rationalistic and sentimentalist approaches, leading to a cognitivistic approach to ethics.

Christian Blum’s *Determining the Common Good: a (Re-)Constructive Critique of the Proceduralist Paradigm* reconstructs the classical proceduralistic paradigm in politics and provides a revision of the paradigm. Arguing that traditional proceduralism cannot explain the citizens’ possibility to be in

¹¹ The opposite point of view is argued by Max Weber, who claims that ethical-normative beliefs have the status of ultimate value axioms (*letzte Wertaxiome*), which enjoy critical immunity and so they are not liable to discussion. See Weber (1917) and see Fittipaldi (2003, 263).

error about the common good, the author proposes to integrate democratic procedural criteria with specific substantive and objective standards of adequacy that should be determined by experts.

Finally, Roberta Sala's *Reasonable Values and the Value of Reasonableness. Reflections on John Rawls' Political Liberalism* criticizes Rawls' proceduralistic account of reasonableness, since it assigns to the "reasonable" only a place in the political debate, while the "unreasonable" are expected to become reasonable or to be paid control. The author maintains that, in order to realistically deal with pluralism, political liberalism should open public debate to those persons who are not "reasonable" – they offer reasons that are neither public nor shared by other reasonable citizens – but who do not represent a danger for the just society.

Given the above phenomenological account on the rationality of norms, since they are understood as liable to be related or not related to truth (related or not related to true or false statements about values), it seems to us that Rawls' Theory of Reasonability risks cutting the source of the rationality of the norms, their possible relation to truth. Doing this, his position could be liable to become in practice tyrannical in the exclusion of non-reasonable opinions (according to which criteria?) and in theory relativistic, denying to the plurality of public opinions the dignity of their pretension to be true. Otherwise, his move could be understood as the request to public actors to switch from the modality of certainty in the pretention of truth, to the modality of plausibility, from rationality to reasonability, in order to avoid fundamentalisms in public debates. Sala pleads for paying attention also to the opinions that refuse this pragmatic attitude, motivated by liberal values to not exclude, but to include non-liberal opinions in the public debate.

Finally, we are very happy to present in the last, special section of the current issue two interviews that were conducted by the editors with Martha Craven Nussbaum and by Valentina Bambini, Cristiano Chesi and Andrea Moro with Noam Chomsky.

In her interview, Nussbaum gives us some insight into her current research project, that will be published as "Political Emotions: Why Love Matters for Justice" by Harvard University Press in 2013. Moving to her liberal account, she stresses the role of emotions and liberal arts in the educations of the citizens of a decent liberal society. Rationality and reasonability are not enough in her account, since every person has to learn to master her passions and to educate her sensitivity in order to have cognitive access to social values. This project is linked with the general Nussbaum's attempt to discern human capabilities in order to formulate the fundamental constitutional principles, that are always liable to be reviewed and improved. Here the role of emotions and liberal arts stands out: they refine human sensitivity and allow us to have cognitive access to renewable interests and claims, playing a role in their recognition through rights, in the limits of what is ought by each to everyone. The interview was held in Cologne, at the time of her Albertus Magnus Lectures 2012 (June 19th-21st). For this opportunity, we would like to thank the *a.r.t.e.s. Graduate School for the Humanities Cologne* (<http://artes.phil-fak.uni-koeln.de/>), which is an international partner of our Research Centre PERSONA.

In his seminar conversation Chomsky starts from the mathematical properties of language and discusses theoretical and epistemological consequences of the research on the biological foundation of language. He also poses new questions that such rational inquiry opens up and that can possibly get an answer in a future. The seminar was held in Pavia, on occasion of Chomsky's visit to the Institute of Advanced Study (IUSS) on September 15th, 2012, and was based on questions proposed by graduate and undergraduate students and reorganized by the staff of the IUSS Center for Neurolinguistics and Theoretical Syntax (NeTS). We thank for this the NeTS and the IUSS (<http://www.nets.iusspavia.it/>), and we welcome them as scientific partner of our Research Centre PERSONA.

4 Interviews

4.1 Martha Nussbaum on Political Emotions

4.2 Noam Chomsky on the Biological Foundation of Language

REFERENCES:

- Berger, P. L. and Luckmann, T. (1966), *The Social Construction of Reality: A Treatise in the Sociology of Knowledge*, Anchor Books, Garden City (New York), 1966;
- Bobbio, N. (1948), *Introduzione alla filosofia del diritto*, Torino, Giappichelli, 1948;
- Conte, A. G. (2001), “Dimensions of Nomic Freedom”, in I. Carter, M. Ricciardi (eds.), *Freedom, Power, and Political Morality. Essays for Felix Oppenheim*, Palgrave, New York, 2001, pp. 69-78;
- Conte, A. G. (2004), “Kanon. Filosofia della regolarità”, *Sociologia del diritto*, 3, 2004, pp. 5-22;
- Conte, A. G. (2007), “Regola eidetico-costitutiva vs. regola anankastico-costitutiva”, in L. Passerini Glazel (ed.), *Ricerche di filosofia del diritto*, Giappichelli, Torino, 2007, pp. 48-68;
- Conte, A. G. (2011), “Tetracotomia della regolarità”, in Amedeo G. Conte, *Sociologia filosofica del diritto*, Giappichelli, Torino, 2011, pp. 73-86;
- De Vecchi, F. (2012), “Making the Social World, or on making our Everyday life World”, *Phenomenology and Mind* 3, 2012, pp. 14-21;
- Fittipaldi, E. (2003), *Scienza del diritto e razionalismo critico. Il programma epistemologico di Hans Albert per la scienza e la sociologia del diritto*, Giuffrè, Milano, 2003;
- Garfinkel, H. (1967), *Studies in Ethnomethodology*, Englewood Cliffs, New Jersey, 1967;
- Hartmann, N. (1933), *Das Problem des geistigen Seins. Untersuchungen zur Grundlegung der Geschichtsphilosophie und der Geisteswissenschaften*, De Gruyter, Berlin, 1933;
- Hartmann, N. (1925), *Ethik*, De Gruyter, Berlin-Leipzig, 1925;
- Hersch, Jeanne (1990), “Le droits de l’homme d’un point de vue philosophique”, in J. Hersch, *L’exigence absolue de la liberté. Textes sur le droits humains, 1973-1995*, Genève, MetisPresses, 2008, pp. 103-125;
- Husserl, E. (1950), (Hua), *Husserliana: Edmund Husserl – Gesammelte Werke*, Springer, Dordrecht, 1950ff.;
- Luckmann, T. and Schütz, A. (1975), *Die Strukturen der Lebenswelt*, Luchterhand, Darmstadt und Neuwied, 1975 (New Edition: Suhrkamp, Frankfurt am Main, 1979);
- Mulligan, K. (2004), “Husserl on the ‘logics’ of valuing, values and norms”, in B. Centi & G. Gigliotti (eds.), *Fenomenologia della Ragion Pratica. L’Etica di Edmund Husserl*, Bibliopolis, Napoli, pp. 177-225;
- Nam-Im-Lee (2006), *The future of Phenomenologica Sociology*, <http://www.newschool.edu/nssr/husserl/Future/Part%20One/PartOneFrames/PartOne.html>;
- Roversi, C. (2012), “Mimetic Constitutive Rules”, *Phenomenology and Mind* 3, 2012, pp. 144-151;
- Scheler, M. (1913-1916), *Der Formalismus in der Ethik und die materiale Wertethik. Neuer Versuch der Grundlegung eines ethischen Personalismus*, in *Gesammelte Werke*, vol. II (1954, 2000);
- Searle, J. R. (1969), *Speech Acts. An essay in the Philosophy of Language*, Cambridge University Press, London, 1969;
- Searle, J. R. (1995), *The Construction of Social Reality*, Penguin Press, Harmondsworth, 1995;
- Searle, J. R. (2010), *Making the social world. The structure of human civilization*, Oxford University Press, Oxford, 2010;
- Smith, B. and Zaibert, L. (2007), “The Varieties of Normativity: An Essay on Social Ontology”, in S. L. Tsohatzidis (ed.), *Intentional Acts and Institutional Facts: Essays on John Searle’s Social Ontology*, Springer, Dordrecht, 2007, pp. 155-174;
- Spiegelberg, H. (1933), *Sinn und Recht der Begründung in der axiologischen und praktischen Philosophie*, in E. Heller, F. Loew (eds.), *Neue Münchener philosophische Abhandlungen*, Barth, Leipzig, 1933, pp. 100-142;
- Spiegelberg, H. (1937), *Sollen und Dürfen. Philosophische Grundlagen der ethischen Rechte und Pflichten*, Reidel, Dordrecht, 1989;
- Spiegelberg, H. (1939), “Justice Presupposes Natural Law”, *Ethics*, 49, 3, 1939, pp. 343-348;
- Spiegelberg, H. (1947), “What Makes Good Things Good? An Inquiry into the Grounds of Value”, *Philosophy And Phenomenological Research*, 7, 4, 1947, pp. 578-611;

- Stein, E. (1922), "Beiträge zur philosophischen Begründung der Psychologie und der Geisteswissenschaften", in E. Husserl (ed.), *Jahrbuch für Philosophie und phänomenologische Forschung*, Jg. 5, Niemeyer Verlag, Halle, pp. 1-284. New Edition: Stein, E., *Beiträge zur philosophischen Begründung der Psychologie und der Geisteswissenschaften* (edited by B. Beckmann-Zöller), in E. Stein, *Edith Stein Gesamtausgabe*, Vol. 6, Herder, Freiburg 2010;
- Terravecchia, G. P. (2012), "Emergent Rules and Social Reality", *Phenomenology and Mind* 3, 2012, pp. 166-173;
- Weber, M. (1917), *Der Sinn der "Wertfreiheit" der soziologischen und ökonomischen Wissenschaftlichen*, in *Gesammelte Aufsätze zur Wissenschaftslehre*, Mohr, 1988;
- Znamierowski, C. (1924), *Podstawowe pojęcia teorii prawa. I. Układ prawny i norma prawna*, Fiszer i Majewski, Poznań, 1924. Italian partial translation of G. Lorini: "Atti tetici e norme costruttive", in A. G. Conte, P. Di Lucia, L. Ferrajoli, M. Jori (eds.), *Filosofia del diritto*, Raffaello Cortina, Milano, 2002, pp. 75-80 (new edition forthcoming).

SESSION

1

SESSION 1

NON-INSTITUTIONAL LIFE

Andrea Zhok (Università degli Studi di Milano)
A Phenomenological Discussion of Early Imitation

Marco Spina (Université Paris-Sorbonne)
Norm and Normality, Starting from Merleau-Ponty

Joseph Kisolo-Ssonko (University of Sheffield)
Love, Plural Subjects & Normative Constraint

Neil W. Williams (University of Sheffield)
Against Atomic Individualism in Plural Subject Theory

Alessandro Salice (Universität Wien)
Social Ontology and Immanent Realism

ANDREA ZHOK

Università degli Studi di Milano

andrea.zhok@unimi.it

A PHENOMENOLOGICAL DISCUSSION OF EARLY IMITATION

abstract

The following pages are devoted to an attempt to examine Andrew Meltzoff's discoveries concerning neonatal imitation in the light of Husserl's discussion of Fremderfahrung. We criticise Meltzoff's explanatory model AIM (Active Intermodal Mapping), which is introduced to account for his empirical findings, for two main reasons. First, the AIM model does not seem to properly reconcile the vindication of the intermodal character of imitation with the idea that early imitation is based on organ identification: these two claims seem to be reconcilable only at the cost of sacrificing the active, non reflex-like character of imitation. Secondly, the account of AIM does not fit in with the ordinary first-person experience of adult imitation. In its stead we propose a different explanatory approach, which is consistent with a basic phenomenology of imitation and does not depend on organ identification, but on the "rhythmic resonance" of gestures.

keywords

Husserl, Meltzoff, imitation, cross-modal identity, rhythm

The question of intersubjectivity in Husserl's thought is one of the most important and controversial issues in phenomenological debate. At the same time, the question of intersubjectivity, with special reference to empathy and imitation, has been the object of cognitive science analyses, some of which have also received the attention of the general public. Our aim in the following pages will be to provide a discussion of a well-known scientific finding in the light of, and in sight of, a phenomenological approach to the question of intersubjectivity. There are remarkable methodological difficulties in allowing the naturalism embraced by cognitive sciences and the fierce antinaturalism of Husserlian phenomenology to interact. We will not try to properly settle here this intricate question, but we just intend the following analysis as a tentative exemplification of a phenomenological use of experimental results in psychology. More specifically, we will briefly re-interpret Andrew Meltzoff's experimental discoveries concerning neonatal imitation by providing them with a phenomenologically inspired interpretive framework, alternative to the one proposed by Meltzoff.

Husserl devotes great attention and thousands of manuscript pages to the problem of the foundation of intersubjectivity, and we cannot try to provide here any plausible resume of such formidable analysis. We will limit ourselves to recall the special role played by the question of *Fremderfahrung* (the experience of the Other). In one of the crucial theoretical steps of his analysis, Husserl tries to explore our primal intuition of the Other, understood as an *alter ego*, an other-subject-like-me. The leading question is: how can we recognise the Other, so that we can discern at the same time the Other's extraneousness and the Other's identity with the Ego that we ourselves are? According to Husserl we cannot have an immediate *primordial* access to the *alter ego*, because this would not allow for our intuition of the Other's alterity: we could not distinguish the Other from ourselves. Husserl tries to explain our intuition of the Other by resorting to the idea of an *apperception* supported by an *analogy* between our body and the Other's one (Hua I, 140-141). Apperception in Husserlian terms is an *association* which pre-delineates the completing traits of the perceived entity. Apperception is an association of *essential* and not merely psychological character. This means, among other things, that the "analogy" between our body and the Other's body cannot be traced back to a contingent external "resemblance". Husserl does claim that there is a kind of resemblance at the roots of the constitution of the "pairing" (*Paarung*) that supports our identification of the Other as such, but the nature of such resemblance is clearer as to what it is

not, than with regard to its positive traits. It is commonly acknowledged that Husserl's treatment of the issue of *Fremderfahrung* remains partially unfulfilled. Our present purpose is to see if it is possible to draw from known empirical investigations some suggestions relevant to that Husserlian problem, while preserving the methodological constraints that phenomenology requires.

Over the last thirty years Andrew Meltzoff and colleagues have produced a series of pathbreaking studies, which have deeply influenced our understanding of children's imitation. Meltzoff's experiments have shown that a sort of imitative process, which we will call *proto-imitation*, can be found even in neonates few hours after birth; such imitation concerns also gestures, like tongue protrusion, which do not allow any visual coupling of the gesture to be imitated with the imitating one (in the absence of mirrors).

The most crucial experiments devised by Meltzoff *et al.* (1977; 1979; 1983) show what follows: after the display of some gestures (tongue protrusion, lips protrusion, mouth opening, and, to older infants, finger movements and side movements of the head) a *majority* of newborns were showing a *greater* amount of repetitions of the displayed gesture than of other gestures. For instance, in an experiment where the experimenter alternately performed sessions of tongue protrusions (TP) and mouth openings (MO), the infants were performing more often MO than TP in the MO session (on average 7.1 vs. 5.4), and they were more often performing TP than MO in the TP session (9.9 vs. 6.5) (Meltzoff and Moore 1983, 705). Also the duration of the infant's gestures was correspondingly greater for the displayed ones than for the others (*ibid.* 706). These results may appear weak, because of their merely statistical consistency, but thenceforth they have been repeated many times, by different researchers and with further methodological caveats. It is important to note that the displayed gestures (TP and MO) belong to the ones that newborn infants *spontaneously* perform. Therefore, the relevant imitation does *not* produce the institution *ex novo* of an unfamiliar gesture after an exemplification. What happens is rather conceivable as the summoning of a *propriocepted* gesture by means of a *visually perceived* gesture. These experimental results must be qualified by the following limitations:

1. Most infants are always producing *some* (even wholly incongruous) responses and, in a minority of cases, even responses akin to the alternative exemplification (e.g., TP responses to MO examples).
2. The responses of the newborns are not immediately emerging in an accomplished fashion, but are gradually approximating a clearer and clearer replication over the course of the experiment.
3. Many infants (more than 60%) do not respond at all or interrupt their responsive involvement, because of greater urges, in the course of the experiment.

As Meltzoff notices, the looseness in responding (1 and 3) and the apparent dependence on learning (2) are incompatible with responses governed by fixed action patterns or reflexes. Responses do not appear to be "ballistically" triggered from appropriate stimuli. Unlike reflexes like the palmar grasp reflex and the sucking one, the working of proto-imitation rather seems to show the character of a primitive *response of "sociability"*, requiring a relaxed situation and an *exploratory attitude*.

At this point Meltzoff tries to provide a causal explanatory account of the mentioned experimental outcomes. He does so by resorting to an explanatory model called *Active Intermodal Mapping (AIM)*, whose general traits are the following:

[I]mitation is a matching-to-target process. The active nature of the matching process is captured by the proprioceptive feedback loop. The loop allows infants' motor performance to be evaluated against the seen

target and serves as a basis for correction. According to this view, the perceived and produced human acts are coded within a common (supramodal) framework which enables infants to detect equivalences between their own acts and ones they see. (Meltzoff and Moore 1997, 180)

These general traits of the AIM seem to be required by the descriptive basis of the experiment: the idea that proto-imitation is a “matching-to-target” process is supported by the fact that apparently proto-imitative acts progressively approximate to the exemplified ones. And the supramodal dimension to which proto-imitative acts must refer is required by the fact that apparently imitation can take place between gestures belonging to different sense modalities (e.g. vision and proprioception). As we will see, this part of Meltzoff’s proposal is consistent with phenomenological descriptions. But Meltzoff goes beyond this explanatory level. He tries to discern the “core mechanism” of imitation by postulating a system of “organ identification” active at birth. The idea is that the newborn is able to identify her own body parts as “corresponding” to the Other’s parts and limbs. This idea is introduced by Meltzoff in order to explain both the general transmodal ability to imitate and the appearance that babies sometimes start the imitating process just by activating the relevant body part (e.g. unspecific activation of the tongue in imitation of TP) (Meltzoff and Moore 1997, 183). And how is supposed such organ identification to take place? Meltzoff formulates two hypotheses. The first one is that organs are innately identified on the basis of their *form*, which we are evolutionarily predisposed to recognise. This hypothesis is suggested by the apparently analogous discovery that neonates are innately sensitive to human faces, that is, to general face-like configurations. This option is the one that Meltzoff prefers and is also an interpretation that has been later endorsed by Shaun Gallagher (2005) in support of his idea of “body schema”. An alternative option mentioned by Meltzoff is that organs could be identified through their unique spatiotemporal pattern of movement (“kinetic signature”): the ways in which tongues, fingers or arms respectively move have specific kinematic constraints that may be sufficient to identify each organ.

At this point, we have to observe in passing that the explanatory proposal that Meltzoff embraces (the first hypothesis) departs from any possible consistency with phenomenological analyses. The reason for this departure is the naturalistic assumption that in the last instance explanations must take the form of a reduction to spatiotemporal causes. Meltzoff rightly remarks that imitation must take place at a supramodal level (that is, at a level common to different sensuous modalities), but then he takes for granted that the roots of both modal and supramodal levels must be causally traced back to events in physical space (brain mechanisms). Husserlian phenomenology would object that spatial determinations as such are *not modally innocent*, since they are *constituted* with reference to specific modal acts (visual and tactile kinestheses). Insofar as we assume that spatiality and causality are beyond the scope of experiential constitution, we just enrol in the mainstream naturalistic club.

But this observation is wholly internal to the phenomenological horizon and is destined to sound unconvincing to naturalistically oriented ears. Let us therefore try to show the limits of Meltzoff’s explanatory proposal from within. Let us take his first option: here imitation as a matching-to-target should be guided by organ identification, which in his view implies the identification of *forms* and *relations between forms*. He calls such *innately cued responses* “organ relation (OR) end states”: they are the propriocepted innate responses that correlate with visual configurations like tongue-to-lips, tongue-between lips, etc. (Meltzoff and Moore 1997, 184). Here however we have a problem: how can the identification of seen and propriocepted *organs* (or OR end states) be conceived if they are all to be considered “*forms*”? Such “*forms*” cannot be *visual* (spatial) forms, as the mentioned case of the neonatal recognition of faces might misleadingly suggest. But if they are not forms in a visual sense, how else should we understand their formal nature? If we look for *supramodal* roots of

protoimitation, the model of the inborn sensitivity to facial traits is inapplicable. Such forms cannot unilaterally belong to any single *modal* dimension.

But could we not just say that some visual forms are simply able to prompt appropriate propriocepted responses? Could we not just suppose that unknown evolutionary developments led our organism to produce appropriate couplings between different modal percepts like seen OR end states and propriocepted OR end states, and that there is nothing more to ask? Of course we could, but it is not what we need either to account for Meltzoff's findings or to account for the mature forms of imitation that we experience in the first person. Inborn immediate coupling of responses is pertinent to physiological reflexes, but could not support the *active* and *generative* character of imitation. There is no doubt, as Husserl has often argued, that we rely on countless passive responses that are not in the power of our conscious activity, but ordinary imitation is an activity guided by an attempt to approximate a target, and *such target cannot be the particular sensuous exemplification that we have in front of our eyes*. It must be a supramodal or transmodal type. Strangely enough, although Meltzoff is perfectly aware of the necessity to resort to a dimension of active representation in order to account for imitation, he does not see that an innate coupling of organs (or OR end states) would never provide such a dimension: even if the visual appearance of eyes, tongue or lips would elicit immediate activation of eyes, tongue or lips, this could not account for the imitation of gestures. Inborn organ identification of a kind may be an initial cue that eases early imitation, but cannot explain imitation and much the less can support the recognition of Otherness. In fact, it is precisely the focus on organ identification that seems to be misplaced. While it is well possible that some reflex responses contribute to an early bodily orientation in *some* newborns (recall that the majority of babies subjected to the experiments do not provide results, either positive or negative), immediate matching of seen and felt organs (or organ relations) cannot support the imitation of gestures.

Meltzoff's idea seems to be that a series of identifications of relative position between organs (OR end states) could work somehow like dots to be connected by pencil: if I proprioceptively know what corresponds to visual tongue-in-mouth and tongue-between-lips, then I could perform imitation as drawing a line between such end points. But this idea, in the absence of essential supplements, would be clearly inadequate: if the relevant "dots" are reflex responses, they *do not take place in any unitary representational space*, as dots on paper would do. This means that we have no guide telling us how to "fill the gap" between reflex responses (which, by assumption, display correlations between visual events and propriocepted ones), because such responses just do not belong to a common representational space: in the experiential "content" of my eye-blinking and of my knee-jerk respectively there is nothing that posits them in specific *spatial* relations, which can be attributed only by reflecting on my body image. Meltzoff, it must be noted, is apparently aware of this problem and thinks that such "connective tissue" between innately cued spots could be originally provided by "body babbling", by which he means the spontaneous repetitive motion of limbs and facial organs that babies produce even before birth. Body babbling should enable the subject to learn how to connect the propriocepted innate responses ("OR end states") with each other and with spontaneous muscular activity so that a kind of body map obtains. This reasonable hypothesis, however, shifts the weight of the argument from the role of innate responses to that of the proprioception of spontaneous muscular activity ("kinaestheses", Husserl would say).

Indeed, if the problem is the one of the transition from the perception of the Other's body to the proprioceptive awareness of my body, the idea of an innate mechanism producing a static mapping of visual positions to propriocepted positions does not do. This model would give priority to the replication of static configurations, which in Meltzoff's own experiments appear to be the exception, rather than the norm. But this model appears to be especially awkward if we consider our first-person experience of imitation, which is after all the ground from which we draw the intuition that neonates may be producing proto-imitative acts.

What happens when we, as adults, try to imitate a facial expression? Except for professionals who may train in front of a mirror, the majority of people are able to immediately grasp a sort of “overall style” of the expression and to reproduce it, even if they do not have any idea about how their face looks. Actually, if our imitative attempts are shown to us (they were filmed, say) we are often baffled by how our face looks. The same could be seen in the imitation of dance steps: uneasiness aside, we can easily imitate the general style of the dance, the rhythm of moves, their lightness or tension, etc. without any specific awareness of the precise position of our feet, of the lines that we are following, etc. True enough, this primary level of imitation can lead to outcomes sadly distant from the model, as any amateur dancer can recognise with the merciless help of a mirror; nevertheless something essential *has been grasped* and can be *recognised* by others. It is important to observe that this stylistic or rhythmic factor can be recognised and transposed even *in the absence of any organ identification*. We can imitate even the joyful attitude of a tail-wagging dog, although no obvious matching of organs is available. We can imitate human, animal and even inanimate motions with limbs that do not belong to the model to be imitated. And our tentative dancing movements can spontaneously coordinate with musical features in an instance of supramodal “imitation”. It seems that there is an imbalance between the spontaneity with which mature imitation can relate to the rhythm/style of perceived motion and the difficulty in picturing to ourselves where our bodily parts and what their spatial relations are. In the following this “rhythmic” or “stylistic” dimension will be named *rhythmòs*, after the ancient Greek term, which has an appropriately comprehensive meaning including ordinary rhythm, proportion and style of patterns; (for an extensive discussion on *rhythmòs* see Zhok 2012, 123 et seq.).

Coming back to Meltzoff’s hypotheses, we should ask whether the second hypothesis that he formulates, without subscribing to it, is more satisfactory. This option implies that organ identification would be obtained by detecting not *forms* but “*kinetic signatures*”. This idea relies on the discovery of the peculiar human sensitivity to biological motions, whose multifarious features are recognised with spontaneity and surprising swiftness (Johansson 1973). This idea, however, does not seem to be adequately exploited by Meltzoff, who is pursuing organ identification as a necessary step towards early imitation. Indeed, it is conceivable that the “kinetic signature” of some gestures is characteristic enough to lead to the supramodal recognition of specific body parts. This is especially possible if we think that newborn infants have a limited repertoire of propriocepted gestures in their command and therefore do not need to perform many refined distinctions in order to find correspondences between seen and propriocepted gestures. The pre- and post-natal body babbling involves few stereotypies and this means that the child does not need a high level of discrimination in order to discern, for instance, MO from TP. Two problems, however, remain: first, this would not amount to *organ* identification, but to *gesture* identification (with incidental identification of the relevant organs), and second, if the correspondence is still read as an immediate elicitation of specific organ activation, then we remain closer to reflex responses rather than to active imitation. Nevertheless, the analysis of early imitation in terms of a kinetic correspondence between gestures opens up a new and more promising view on imitation and its tie to empathy.

It must be noted that the “kinetic identity” between perceived and propriocepted gesture must not properly concern movement as it is ordinarily understood. Although in the third person (i.e., for objective representation) the gestures that I perform and experience in the first person can be described as motions (i.e. changes of position over time), they are not propriocepted by me as something taking place in objective space and time. This means that the propriocepted gesture does not have ordinary “motion” in common with the perceived gesture. But then what do they have in common? We must take seriously the demand that the dimension where imitative matching can take place be *supramodal*. The representational space where motions are placed is inescapably tied to *visual* aspects (motions are represented as events in an ideal visual field), whereas our disposition to “feel in resonance” with seen

gestures does not belong to such representational space: we can spontaneously perceive the “similarity” between the lowering pitch of a whistle and the approaching trajectory of a falling object. Such similarity is no exterior congruence of sensuous traits. Seen motions, propriocepted muscular activations and unfolding sounds can have in common a “way of changing”, an instance of *rythmòs*, to which we must not attribute any particular sensuous aspect: it does not occupy either a visual space or a tactile space or any other sensuously qualified dimension, since its “substance” is rather a *temporal form*. It must be noted that a truly supramodal dimension is necessarily superindividual: it must be possibly valid for an indefinite plurality of sensuous instantiations. Furthermore, it should be stressed that *rythmòs* must also be a *motivating* dimension, in the sense that instances of *rythmòs* are apparently *salient* for the newborn perceiver, whose corresponding gestures are called forth.

A notion like a supramodal motivating *rythmòs* can account for the early imitative response better than any *organ identification*, which can be either derived from the correspondence between instances of *rythmòs* (as “kinetic signatures”), or attributed to inborn reflex-like cues inessential to the imitative process. From this point of view some of the results of Meltzoff’s experiments become more easily interpretable.

As we said above, many newborn babies did provide apparent responses to the exemplified gestures, even if they were the “wrong” answers. But, if we take organ identification to be explanatorily prior and dynamic gesture correspondence to be a secondary achievement derived from the former, then responses that get the pertinent organs wrong (TP instead of MO) are not exactly what one would expect. On the contrary, if we reckon that at first babies respond to instances of *rythmòs* as such (primarily conceivable as biological motion), we can account for the “statistic” trend toward proper matching of gestures. The majority of babies respond just to *instances of rythmòs*, and the more the corresponding process unfolds, the better some of them refine the dynamic pattern of their responses, down to the appearance of imitation proper.

The character of motivating salience of the instances of *rythmòs* accounts also much better for the paradoxical nature of imitation and empathy, such that we experience ourselves at the same time as identical with and different from the Other. What is proximately felt, what is “mine”, is *rythmòs*, whereas its contingent modal realisations are irreducibly “other”, extraneous. What is in common is what enables communication, primarily emotional communication through the style of gestures (softness, regularity, abruptness, violence, etc. are immediately detectable) (see Stern 2000, 48 et seq.). What remains ineradicably different is the specific sensuous implementation of the gestures as well as the relevant character of reactivity or spontaneity of each implementation.

Finally, the notion of *rythmòs* is not a mere hypothesis consistent with, but superimposed to, phenomenological observations: *rythmòs* is intuitively available in a plurality of phenomena which go well beyond (early or mature) imitation. The idea that *imitation* could be the primary access leading to empathy with the Other and her “state of minds” seems now somewhat misleading. Our primal access to the Other does not go either through direct coupling to the Other’s body parts, or through direct access to the Other’s feelings (in this case we would *be* the Other). We have direct access to the Other’s *expressions* of feeling, which are embodied in specific instances of *rythmòs*. In other terms, we need not think that we first produce a more or less faithful copy of the Other’s behaviour, on the basis of a structural correspondence of bodies, and then gain a mediate access to the meaning of the Other’s gestures. *Rythmòs* in the gestures is primarily evident and is expressively intelligible even in the absence of the recognition of specific body correspondences: this is the reason why we need no mediate introduction to grasp the expressive motions of animated objects in cartoons (brooms, cars, haunted houses, etc.).

The way of moving is enough to define the object's "emotional dispositions" as well as its targets and spatial orientation, even in the absence of any clear reference to eyes or limbs. This perspective could be critically applied to theoretical views like the one recently proposed by Gallagher (2005), where a body schema endowed with the same structure of the average body image is used as explanatory key for cross-modal and intersubjective relations (see Zhok 2012, 202 et seq.).

REFERENCES

- Gallagher, S. (2005), *How the Body Shapes the Mind*, Oxford: Oxford University Press, 2005;
- Husserl, E. (1973), *Cartesianische Meditationen und Pariser Vorträge*, Husserliana I, S. Strasser (ed.), Martinus Nijhoff, Den Haag (quoted as *Hua I*), 1973;
- Johansson, G. (1973), "Visual perception of biological motion and a model for its analysis", *Perception and Psychophysics*, 14, 1973, pp. 201-211;
- Meltzoff, A. N. and Moore, M. K. (1977), "Imitation of Facial and Manual Gestures by Human Neonates", *Science*, 198, 1977, pp. 75-78;
- Meltzoff, A.N. and Borton R. W. (1979), "Intermodal Matching by Human Neonates", *Nature*, 282, n. 5737, 1979, pp. 403-404;
- Meltzoff, A.N. and Moore M.K. (1983), "Newborn Infants Imitate Adult Facial Gestures", *Child Development*, 54, 1983, pp. 702-709;
- Meltzoff, A.N. and Moore M.K. (1997), "Explaining Facial Imitation: A Model", *Early Development and Parenting*, 6, 1997, pp. 179-192;
- Stern, D. (2000), *The Interpersonal World of the Infant. A View from Psychoanalysis and Developmental Psychology*, New York: Basic Books, 2000;
- Zhok, A. (2012), *La realtà e i suoi sensi. La costituzione fenomenologica della percezione e l'orizzonte del naturalismo*, Pisa: ETS, 2012.

MARCO SPINA

Université Paris-Sorbonne

mar.spina@gmail.com

NORM AND NORMALITY, STARTING FROM MERLEAU-PONTY

abstract

The present work pursues two main objectives. The first is to pinpoint, by means of a series of philosophical arguments, the meaning of the word “norm” and, specifically, to explain the tension between normality and normativity in detail. The second, closely related to the former, hinges on Merleau-Ponty's personality, and aims at putting the “space” of the norm and the key notion of normality in a new form, starting from some essential elements of his writing. This effort is particularly significant in The structure of Behavior, Phenomenology of Perception and in The Primacy of Perception and its Philosophical Consequences.

keywords

Merleau-Ponty, normality, normativity, origin and structure of norm, judgement on normality

1. Meeting the Problem

1.1

Normality and Normativity

The starting point for our dissections is an essay entitled *Normalité et normativité. Entre phénoménologie et structuralisme*, written by Bernhard Waldenfels, one of the most important living exponents of phenomenology. In this essay Waldenfels discusses the question of the norm on the basis of the pair normal/normativity:

le terme “norme” originariamente destiné à dénoter un instrument de mesure, un cordeau ou une équerre dont se sert l’architecte, se divise en “normalité” et “normativité”. Les deux termes se ressemblent fortement, comment peut-on alors expliquer leur différence? Depuis Hume ou Kant, on tend à concevoir la normalité comme un complexe de règles descriptives, et la normativité comme un complexe de règles prescriptives. Les unes portent sur des questions de fait, les autres sur des questions de droit. Le comportement humain est soumis aux deux dimensions: normalité et normativité (Waldenfels 2005)¹.

In Waldenfels’s opinion, such an alleged antinomy between normality and normativity succumbs to two serious failings. The first concerns the incarnation of the norms, the second concerns their genesis. On the one hand, what is omitted is the entrenchment of each norm in the world of life: the norms become incarnated as habits, usages, and customs. What comes out is consequently not assessed as right or wrong, good or bad, functional or not functional, but as convenient or inconvenient. On the other hand, what is left out is the fact that norms cannot exist unless they have been learned and appropriated, and that they therefore originate through a genesis revealing itself, to some extent, as an event of *Stiftung* (foundation), even if the latter settles into a pre-history of which no trace can be found anymore. A similar position is presented by Norberto Bobbio, notably in his formulation of the entry *Judicial Norm* in Bompiani Encyclopedia of 1964, where the Italian philosopher of law describes the relationship between the categories of judicial normativity and of normality as follows:

Del significato comune di “norma”, come del resto di “regola”, sono elementi caratteristici, rispetto alla funzione, il fine prescrittivo (dove viene l’aggettivo “normativo”); rispetto al contenuto, la tipicità del

¹ “The term ‘norm’ originally denotes a measuring instrument, such as the cord or the square that is used by architects. The ‘norm’ is divided in both ‘normality’ and ‘normativity’. These two are very similar. So what are the differences between them? Following Hume and Kant, normality is conceived as a plethora of descriptive rules, and normativity as a complex of prescriptive rules. The descriptive rules lead to questions of fact, whereas the prescribed ones lead to questions of law. Human behavior is subject to two dimensions: normality and normativity”.

comportamento (dove viene l'aggettivo "normale"). Una norma è, si potrebbe dire, una proposizione normativa che tende a riconoscere e stabilire un comportamento normale: il carattere della normatività riguarda il fine, quello della normalità il risultato (Bobbio 1994)².

Even if a distinction is made between normative--all that relates to the purpose of the norm--and normal--all that concerns the content and the effect of the norm--in Bobbio's definition the relationship between the categories of normal and of normative is manifest: every judicial norm "tends to identify and establish a normal behaviour". Therefore, every norm seems to imply a normality, pre-existing and giving legitimacy to the same norm: the normativity of the normal. Likewise, every norm sets the conditions of normality, "normalizing" the receivers' behaviours: the normality of the normative. Nonetheless, a problem arises from this context.

Normality is one of the most difficult-to-define and most polyhedric of concepts, and certainly one among those which have given rise to the most heated debates. It is a tricky and ambiguous notion, surrounded by an aura of social dangerousness, because it can easily spread pressures, inhibitions, and discrimination within a community. The fact that in the deconstructionist and post-modern ambit the pluralistic perspectivism of "discursive formation" and language games has prevailed over "meta-narratives" and over the establishments made of universal and eternal truths, has progressively led to condemnations and attempts at total elimination of the concept of *normality*, meant as an instrument of power and of oppression, of the institutional and established power - the ruling class - and as the different, the dissident, the Other (Foucault 1970; Lyotard 1984).

On the contrary, everyday language has continued to resort to substantives and adjectives referring to *normality*, being evidently legitimated by the existence in the common view of the world of something matching the notion of *normality*, beyond any theorization and ideology. Referring to language as the legitimating foundation of *normality* may seem quite risky because of the imprecision and ambiguities its use reveals. The adjective *normal* and the substantive *norm*, from which the former comes, have, in different contexts and perspectives, completely different meanings. In Italian, for instance, the word *norm* shifts from the most settled use, namely the ethical and juridical one as *rule*, to the medical one as physiological standard meaning, e.g., "to be up to/below standard", to the social one as *usage* and *custom*. The adjective *normal* wavers in a wide semantic range, according to the context in which it is used, as synonymous with *clear*, *logical*, *right*, *habitual*, *coherent*, *customary*, *usual*, *common*, and *natural*. Nonetheless, beneath this multivalence it is possible to trace a common characteristic of the different meanings, all together describing an attitude clearly based on a mutual and shared logic, which occurs almost regularly.

The concept of normality concerns the behavioural relationship between the subject and the environment, or rather between the subject and the world, and in the subject-world polarity the norm that determines normality takes shape.

Therefore, it can be supposed that the problem of normality and of its elusiveness lies concealed in the subject-world dualism that derives from the dualism of the Cartesian subject. As a consequence, it is necessary to reconsider the concept of normality on the basis of the supersession of the Cartesian dualism and on the basis of a notion that reintegrates the subject and the world into an organic whole and redefines them in relation to one another. In considering all of this, light will be cast on some themes of Merleau-Ponty's thought, pointing out their salient characteristics, to find the necessary elements to answer to the question we have hereby raised: what is a *norm*? What is *normality*? In what way can we talk about it?

² "The common understanding of 'norm' and 'rule' is characterized by prescription regarding its function (from where the adjective 'normative' derives), and by the typicality of the behavior regarding its content (from where the adjective 'normal' derives). One could say that 'norm' must be intended as a normative proposition that tends to recognize and establish a normal behavior: normativity is linked to its aim; normality pertains to its result".

1.2
Where is
the Norm?

Leaving aside for the moment the observations made by Merleau-Ponty in *The Structure of Behaviour* and *Phenomenology of Perception*, it is important to focus in the first place on another crucial passage of his work, thus making the questions we are discussing come out even more drastically.

On November 23, 1946, Merleau-Ponty gave a lecture at the *Société française de Philosophie*, wherein he illustrated the outcomes described in *Phenomenology of Perception*, which he had published the year before. The text of this report, together with the transcript of the heated debate it sparked, was later published under the title *The Primacy of Perception and its Philosophical Consequences* (Merleau-Ponty 1964).

The thesis of Merleau-Ponty is firm: “The perceived world is always the presupposed foundation of all rationality, all value, and all existence. This thesis does not destroy either rationality or the absolute. It only tries to bring them down to earth” (Merleau-Ponty 1964, 13).

It is a thesis which is even more relevant now than it was at that time. What is at stake is a kind of philosophy which is able to relate that which is human without relinquishing what makes him human, namely “rationality and the absolute”. This really is a current matter for at least two reasons. On the one hand, because the naturalism of cognitive science and biotechnology has really tried to “bring” the human being “down to earth”, but at the price of stripping him of everything that makes him human. According to classical cognitive science, in fact, thought is an immaterial and abstract calculation, while according to biotechnologies it is some kind of effect resulting from the human genome. In both the cases, it is manifest that no space is left for the human, at least as it has been conceived from the Greek age until now.

On the other hand, the lecture of Merleau-Ponty is useful since it perfectly depicts the problems each materialism has posed: its formulation, in and of itself, shows it clearly when taking on the strange task of “bringing” the human world “down to earth”. Now, such a demand can be put forward only if it is believed that that world, itself, would float and be unrelated to real life. In Merleau-Ponty’s opinion it is precisely so: the human world needs a foundation and this foundation cannot be other than “the perceived world”, the world of body and of its senses. “We can only think the world because we have already experienced it; it is through this experience that we have the idea of being, and it is through this experience that the words ‘rational’ and ‘real’ receive a sense simultaneously” (Merleau-Ponty 1964, 17).

The statements by Merleau-Ponty, anyway, face an apparently insuperable problem: if my body is the basis and the unit of measurement of my experiences, and contrariwise your body is the basis and the unit of measurement of your experiences, how can we succeed in creating a common world where we can meet, discuss, produce science, etc., if science, ethics and everyday experience assume the existence of a common and collective world, the existence of an objective world?

It is exactly what one of the participants in Merleau-Ponty’s lecture, Bréhier, commented on from an ethical point of view: “the other is ‘reciprocable to me’ by reason of a universal norm. Where is your norm?” (Merleau-Ponty 1964, 31). To be able to consider the other as someone having his own rights, as someone whose point of view is as valid as mine, I have to manage to detach myself from my body and to think of my perspective as equal to his within a universal and objective norm, which is neither mine nor his. How does Merleau-Ponty reply to this criticism? Depreciating the universality of the norm in the relativism of the norms: “there is no given universality; there is only a presumptive universality” (Merleau-Ponty 1964, 31).

The fact that the norm can change and sometimes indeed has to change does not mean that a norm cannot be objective; a subjective norm is no more a norm. But if it is objective, how can it find its foundation in a subjective body?

From this question, you can get to the bottom of Merleau-Ponty’s works.

The definition of the subject and of his relation with the world is gradually developed in Merleau-Ponty's works, through many phases which follow each other, overlap, and merge into one other. The first phase is closely related to the influence of K. Goldstein. Goldstein's studies on subjects affected by neuromotor disorders highlight a unitary system of distinction with progressive levels of disintegration of the behaviour within an organism conceived as a living totality (structure) inseparable from its environment. In this perspective, disease is defined as loss of liberty with respect to the environment, impotence in escaping from fortuitous circumstances and from their drives, and, consequently, in projecting oneself into the future.

In *The Structure of Behaviour* Goldstein's themes cross with Husserl's phenomenology and with Gestaltpsychologie. The criticism of Husserl's pure consciousness and absolute subjectivity, based on the irreducibility of body, leads Merleau-Ponty to reformulate the consciousness-world relation starting from corporeality and perception, meant as the initial relationship between consciousness and the world, beyond the antithesis between idealism and empiricism. Merleau-Ponty's analysis goes beyond the subject-world couple to consider behaviour as the initial unitary phenomenon where the subject and the world are continuously grasped one after the other. The Gestalt influence allows him to formulate a definition of behaviour as structure; in other words, as a totality equipped with an internal principle of diversification.

Behaviour as structure varies according to its belonging to the organic or the inorganic world, on the basis of a different normative structure which is somehow its essence. In such an ambit, Merleau-Ponty enunciates two definitions of norm, both referring to the *organic structure*. The first one reads the norm as the essence of the species, namely as a system of personal constants, which express the ideal conditions of relationship to the environment and of activity: an experienced norm which corresponds to one's own being. The second one reads the norm as the biological *a priori* of an individual: "by 'norm' here one does not mean a 'should be' which *world make* it be, it is the simple observation of a preferred attitude statistically more frequent, which gives a new kind of unity to behavior" (Merleau-Ponty 1963, 159).

In *The Structure of Behaviour* the concept of *norm* is meant as the structure by means of which the behavioural constants of an organism appear to an observer. Or, in Goldstein's words, the *norm* is the *structure*; in Husserl's words, the norm is the *style* of behaviour. As a result, it can be asserted that the perception and the observation of a subject "X" have to reveal the structure of his behaviour, that is, his norm. In phenomenological terms, it may be defined as some kind of eidetic reduction: in this sense it can be maintained that the norm of behaviour of "X" is the essence of "X", caught through perception and based on the certainty of the cognitive act, whose actuality is direct. Anyway, it is important to specify that it is not only defining the norm of "X" but the norm of his behaviour, according to a course already followed by Jaspers, Binswanger and Minkowsky.

This means that the norm, far from being an abstract, fixed and immutable behavioural parameter, takes on its sense only in connection with a "ground", that is to say, with an environmental condition, and defines the best adaptation possible of man to the environment and, consequently, his ideal living conditions. Then, the fulfilment of the norm should lead to a satisfying relationship with the environment and to the adoption of the so-called--by Goldstein--"privileged behaviours".

To the extent that human behaviour is characterized by the supremacy of the symbolic form--which organizes and involves the other forms in its dialectic, thus giving them peculiar sense and significance--its particular way of being in the world is essentially marked by transcendence. The norm of human behaviour, in fact, does not consist in merely searching for a stable equilibrium with the environment. Rather it is, indeed, seeking for something which still is not, for some kind of striving for the virtual, for an effort which is at the same time power of denial and transformation of the real, and which is vehiculated by a symbolizing and conceptualising power, going hand-in-hand with the integration of spiritual activity with corporal activity.

2. Norm and Normality, Starting from Merleau-Ponty 2.1 The Structure of the Norm

**2.2
The Temporality
of the Norm**

This initial structural definition of norm, transpiring from the pages of the first work of Merleau-Ponty, finds a better and more extensive elaboration in his *Phenomenology of Perception* (Merleau-Ponty 1962). The phenomenological analysis of perception puts the pre-reflective aspect of the existence into the foreground, thus localizing the latter in a broader behavioural norm including both the dialectic of forms, so the body-soul integration, and the world-consciousness-work dialectic, which, in *The Structure of Behaviour*, is considered as a privileged behaviour of man. As a result, the norm, far from disregarding its structural characteristic, starts out to become the essential structure appropriate to embrace the whole variety of the *Erlebnisse*, on the general ground of *being-in-the-world*. Let us have a look at this more closely.

The common feature of the pathological behaviours examined by Merleau-Ponty is the loss of the categorial function, namely, in a few words, the loss of the symbolic-virtual openness to the world. In this case the subject has shrunk away and positioned himself on a private behavioural norm of a “minority” kind, where the structure of transcendence is blocked or invalidated. This means that the whole “capability of disposing of a past, a present, and a future” is nullified, that is to say, that the transcendence of the empirical subject to the world is stricken, together with all the essential structures relating to it: intentionality, perception, cogito, it’s own pre-history as open and retrievable control over body workability.

Basically, next to the temporality-transcendence, the very essence of the subject has been hit, which leads us to conclude that temporality is the fundamental norm of human behaviour and is defined as temporal *style*, and that normality consists in human behaviour taking place according to the structure of temporality.

The “normal” disposal of temporality means offers of chances, multiplicity of worlds and of horizons, and fluent integration of the past with the future in a present which is the operational grip of a body exactly on the world. The block of temporality thus signifies the destructuring of the present, not a pure and simple blockage of the future, as suggested by Minkowsky, but the sticking of existence to a behavioural arrangement which cannot be defined as normal in any way and whose most essential feature is being apraxic. The normal subject in his deep structure is, as we are attempting to demonstrate, an *I can*, namely an active subject, mainly defined by a working intentionality and a pragmatic openness. In this sense a behaviour being chronically apraxic must be considered far from the style of the normal behaviour. Moreover, further to the disgregation of temporality, a destructuring of personality and a fragmentation of the subject occur.

Assuming the temporal style of existence as norm generates a structural notion of norm, which does not determine the content of the behaviour but its *how*. Set down like so, the norm can govern the individual-environment relationship, in each variation in its own terms, without having a specific content, and can express the sense organising at the same time the life of a species, of a collectivity or of a single man, all of them being structures. Through this perspective, the question of the temporality of the norm has to be coped with, wondering if it is eternal and immutable or historic.

**2.3
Experience and
Judgment of the
Normality**

The norm, devoid of content, invariable in its expressions of the temporality of the subject, on each occasion takes on the content required by the situation. So it is what changes, even though always remaining the same. What varies is the way in which behaviour expresses the norm, this relation defining the normality. The norm is therefore atemporal, it being structure. Nevertheless, since it also conveys the sense of a behaviour, the norm cannot dodge the reference to the fortuitous situation in which this sense expresses itself and the relation such a sense has with the general sense of the world. Normality in this meaning is outlined according to the ambiguous aspect of temporality, as intersection of what is historic and what is atemporal, thus defining two axes of normality: the first (atemporal) is one of individual disposal of temporality; the second (historic) is one of articulation of the sense of the individual behaviour within the general sense of the world, one of the gears of

individual time in collective time. The latter makes us refer to the inter-subjective dimension and to the central question of normality. Every subjective paradigm of normality has to find confirmation in the intersubjective dimension, as in the same way the individual norm cannot be flattened into the environmental norm. The judgement of normality is at the same time judgement of the individual on himself and judgement of the other on the individual, articulated synchronically and diachronically. Particularly on this last topic we can observe that as a matter of fact, following the pathway of Merleau-Ponty more closely, it can be noticed that the supersession of the Cartesian subject based on the notion of body precisely opens at once, according to the author, an intersubjective dimension which leaves out any sort of solipsism. Man finds himself, among other men, as already given in the world, and his perception discovers *typicalities* similar to his ones. Subjects set up, with one another, a working knowledge, where a common way of experiencing the world is expressed. Communication is behavioural and is immediately understood by the perception of the others.

The body of the other, by virtue of its symbolizing power, is significant and reveals itself as a *cultural object*, and it is just for this reason that perception can perform an *intentional transgression* with regard to it, thus avoiding intellectual mediation and thereby winning access to the sphere of the other.

The intersubjective relationship has always lain in an interworld, it being the place of meeting of two ipseities that disclose themselves in a reciprocity of styles, rhythms, and time. Said reciprocity is not an absolute transparency, but an initial possibility of communication and immediate understanding, on which the possibility of speech and dialogue itself is based.

The relation between ego and alter-ego bursts on the scene of the common world, where both have always been involved and have always met, and demands a necessary character. This is a physical and social world, full of collective cultural meanings in which anyone can cooperate in creating and destroying. In this socio-cultural world, which is man's "natural" world, non-material beings can be found. Said beings originate in the interworld as world validities having a universal value and which, belonging to the *they*, go beyond the individual in space and time and preconsciously live with him as *modes of (co)existence*.

In this ground of settled and experienced validities, the subject's evaluative and interpretative approach, from which he derives his parameters and his legitimacy of judgement, has put down its roots. And it is right on this ground that the question of normality has to be brought back and placed, if the ambiguities that usually characterize it are to be extirpated. Every objective definition of normality is given as ambiguous and paradoxical, due to the fact that its parameter is the world's pre-logical *logos*, materialized in the *they* and not in the commensurate evaluations of any science.

The normal is the pre-reflective and normality, thought as pre-reflectivity, cannot be caught by means of the scientific objectifying instruments of thought, but it manifests itself surrounded by an aura of ambiguity in the first step of perception. This perception is not an arbitrary intuition.

One's own personal history as experience, the history of the other as behaviour, and the social history as *typical* of the human behaviour converge in the perception of normality and are grafted onto the validity of the normality as an intersubjectively-determined value.

Perception throws open for us the door onto the existence of the other through his body; what the abnormal individual expresses through his behaviour is his constraint in a private world and his impossibility of communication, whereas his body openly communicates what he would rather not reveal. Who stands in front of us is not the inhabitant of an unknown world, but an exile from our world. His own existence is globally connoted by a kind of suffering which has nothing in common with the ordinary trials we are used to feeling, even thought in their worst forms; it is a radical and incommunicable suffering. Such suffering is incommunicable, but not totally incomprehensible from the outside. The possibility of understanding an abnormal form

of existence is left open by the common world foundation, which binds individuals together, and by the willingness to “get in the game” and take on the “centrifugal” norm of the abnormal behaviour. This possibility of getting in the game does not disconcert my being in the world. An abnormal form of existence cannot put our world at stake completely, since it represents but a minority form. Rather, when we get in the game to understand this form of passing existence, we reassert our possibility of transcending its constraint.

REFERENCES:

- Arendt, H. (1958), *The Human Condition*, The University of Chicago Press, Chicago, 1958;
- Bergeret, J. (1974), *La personnalité normale et pathologique. Les structures mentales, le caractère, les symptômes*, Dunod, Paris, 1974;
- Bobbio, N. (1994), *Contributi ad un dizionario giuridico*, G. Giappichelli, Torino, 1994;
- Canguilhem, G. (1989), *The normal and the pathological*, Zone Books, New York, 1989;
- Heidegger, M. (1962), *Being and Time*, Blackwell, Oxford, 1962;
- Foucault, M. (1970), *The Order of Things. An Archeology of The Human Sciences*, Pantheon Books, New York City, 1970;
- Kelsen, H. (1960), *Pure Theory of Law*, University of California Press, Berkeley, 1970;
- Lyotard, J.F. (1984), *The Postmodern Condition A Report on Knowledge*, Manchester University Press, Manchester, 1984;
- Merleau-Ponty, M. (1962), *Phenomenology of perception*, Routledge, London, 1962;
- Merleau-Ponty, M. (1963), *The Structure of Behavior*, Beacon Press, Boston, 1963;
- Merleau-Ponty, M. (1964), *The Primacy of Perception, And Other Essays*, Northwestern University Press, Evanston, 1964;
- Sartre, J.P. (2007), *Existentialism Is a Humanism*, Yale University Press, New Haven, 2007;
- Shiner, R. (1992), *Norm and Nature*, Clarendon Press, Oxford, 1992;
- Waldenfels, B. (2005), “Normalité et normativité. Entre phénoménologie et structuralisme”, *Revue de métaphysique et de morale*, 1, 45, 2005, pp. 57-67.

JOSEPH KISOLO-SSONKO

University of Sheffield

j.kisolo-ssonko@shef.ac.uk

LOVE, PLURAL SUBJECTS & NORMATIVE CONSTRAINT

abstract

Andrea Westlund's account of love involves lovers becoming Plural Subjects mirroring Margaret Gilbert's Plural Subject Theory. However, while for Gilbert the creation of a plural will involves individuals jointly committing to pool their wills and this joint pool directly normatively constraining those individuals, Westlund, in contrast, sees the creation of a plural will as a ongoing process and she rejects the possibility of such direct normative constraint. This rejection appears to be required to explain the flexibility that allows for a central place for reciprocity in loving relationships. However, this paper argues against the existence of such flexibility and presents instead the case that variance in the normative pain of rebelling against the collective will should be understood by replacing Gilbert's notion of all-or-nothing pooling of wills with an account that sees wills as becoming entangled through levels of identification with the plural subject.

keywords

Constraint, collective intentionality, love, union, plural subject theory

*“Whenever infatuation begins, if given the opportunity
it transforms itself into a continuing romantic love
With this continuing romantic love it feels
to the two people that they have united to form
a new entity in the world, what might be called a we.”
(Nozick 1995, 232)*

1. Love and Forming a We

Two strangers kissing on a hedonistic night out drinking may well remain nothing more than lustful individuals, however on many accounts of love (such as that of Robert Nozick quoted above) moving beyond such initial stages of desire involves becoming united in some real sense; a move from a mere set of separate *Is* into a combined *we*. Marriage might be thought to be the clearest form of such a move, with its formal vows and declarations, however such relationships are possible without legal ties and for want of a name I will dub this kind of relationship a “committed romantic relationship”, for brevity henceforth referred to as a CRR¹. Talk of becoming a *we* may well fit with our experience of love but is it *literally* true? Even when we are most deeply in love we remain distinct biological entities; never actually becoming one creature with two heads. To meet the challenge of making sense of this it seems obvious to look to the growing literature around what has been called “Collective Intentionality”. Of particular relevance to Nozick’s work is the “Plural Subject Theory” (PST) of Margaret Gilbert; for his talk of forming a *we* mirrors her focus on “... the central sense of the pronoun ‘we’” (Gilbert 1992, 152) but I will start with the account of CRRs proposed by Andrea Westlund as, while in some ways similar to Gilbert’s PST account², it poses some illuminating challenges to it.

Westlund suggests that CRRs (or as she calls them relationships of “companion love”) (Westlund 2008, 558) involve the formation of a plural subject with a particular structure; a structure that respects the reciprocity that she sees as essential to love. While Westlund invokes the idea of a plural subject she rejects the claim that normative constraint arises directly from being part of one. She thinks this necessary to account for the flexibility between parties she sees at the heart of loving relationships. This normative constraint is a key element of Gilbert’s theory and I will challenge Westlund’s rejection of it but in doing so I will argue that CRRs *do* demonstrate a related phenomena that might be confused with

¹ I do not want to claim that my description of CRRs *stipulates* the only way people can be in love, rather just that it *describes* a kind of loving relationship we commonly form.

² In her paper on marriage (2008) Westlund does not explicitly state that she is attempting to apply a Gilbert-esk framework, indeed she makes reference to an number of other authors (Bratman, 1999, Roth 2004, Searle, 1990, Tuomela & Miller, 1988, and Velleman, 1997) and seems to imply that her account is relatively neutral between them. However, the way in which she invokes the notion of “pooling of agency” and sets out the notion of a “shared practical perspective” *does* appear to be particularly fitting with Gilbert’s framework. Her engagement with Gilbert is more explicit in an earlier draft of her paper (published as a “working paper”, Westlund, 2005). She does not appear to have read the work where Gilbert most explicitly deals with love (Gilbert, 1996) though she does appear to be acquainted with a great deal of Gilbert’s work.

Westlunds flexibility i.e. variability of the pain of resisting such constraint. My tentative solution will be to urge the replacing of Gilbert's discrete (i.e. all-or-nothing) voluntary pooling of the wills with a continuous (i.e. open to various degrees) notion of wills becoming entangled through levels of necessary identification with the plural subject.

Westlund worries that the union view of love inasmuch as it seems to suggest that lovers become psychologically or ontologically melded appears to be contrary to something fundamental to her understanding of CRRs: reciprocity. Reciprocity seems to require each agent to value the other's interests. The union view appears to dissolve the interests of each into the *we*. What we can call the "*problem of reciprocity*" is the worry that it is not possible to value the interests of the other if they have been dissolved into a *we*. Rather than give up on the union view Westlund believes this problem can be overcome by developing an account of the union in question that does not entail melding. She takes this to be possible if the union is of the form (*à la* Gilbert) of the creation of a *plural subject*.

For Gilbert plural subjects are the willed creations of individual agents; they are the result of those agents jointly committing to, within a certain scope of activity, *pool their agency*. Imagine a couple taking a trip together, for Gilbert what makes their travelling *together* stronger than merely travelling in proximity (temporal and/or spatial) is that the agents can be said to have jointly manifested and accepted "... willingness to constitute with the other a plural subject with the goal that they travel in each others company" (Gilbert 1992, 163). Plural subject-hood thus involves agents remaining as individuals, i.e. merely placing some of their agency into a shared pool, a pool with a specific goal or stance, rather than becoming psychologically or ontologically melded. Westlund is not entirely clear on how introducing the notion of plural subjectivity helps but I believe we can read her as saying that it avoids the problem of reciprocity because forming a plural subject need not destroy the self and thus the interests of each individual remain intact.

Plural agents can be transitory; e.g. the travellers in Gilbert's example may have met for the first time at the train station and travel together solely for one trip, or they can have more open-ended/extended scopes; e.g. the travellers may be regular fellow commuters. Clearly CRRs are of the latter type, indeed in marriage the participants pledge to be as one *until death parts them*. So far so good, but we might ask, given that plural subjects can have all kinds of implementations, what is the specific scope of a committed *romantic* relationship? Gilbert suggests that it might be "... a specific primary goal: something like the well-being of both parties equally" (Gilbert 1992). While Westlund sees reciprocity as a vital part of what it is to be in a CRR she does not think that it is the direct goal of CRRs rather she claims that it characterises the nature of *how* CRRs try to achieve their main goal, which is simply *living as one*. Living as one will involve all sorts of immediate shared goals; to go for a walk, to buy a sofa, to discipline their child etc. but always "... the projects and plans undertaken by companion lovers are ways of realizing an overarching desire to be (and do) together" (Westlund 2008, 558).

For Westlund achieving the goal of living together *in a loving way* has two elements; firstly, sharing reasons as well as sharing ends, and, secondly, sharing these in a way that engages with the individual interests of each party. Sharing in ends means together aiming at doing particular things *as a couple*. Sharing in reasons, means together taking certain considerations to count in favour for them *as a couple* to do those things. Together these can be thought of as a process of forging a *plural will*. Take a couple who are planning to go on holiday; on Westlund's account this can be contrasted with strangers who, having won a game show, are negotiating a single location

2. Westlund's Plural Subject Account

to be whisked away to – the important difference is that the couple’s aim is not just to come to *independently* converge on a single destination, rather, their aim is to make the decision jointly. So, rather than weighing up all their individual preferences and trying to find a holiday that merely maximises these as far as possible, the couple will try to find reasons that they can agree count in favour of a holiday for them *jointly*; e.g. “I love culture and my partner loves beer but we want to go to Majorca because we want to go somewhere hot”³. All of which is not to say that joint considerations about what counts as a reason for them jointly cannot take each individuals preferences into account, just that what counts as their collective preference is not automatically given by finding the point of maximum cross over between individual preferences. Thus discussion ends not with the discovery of shared reasons but rather in their creation through *joint acceptance* (Westlund 2009, 6). In this sense the plural will is a “... joint product of collective deliberative agency” (Westlund 2008, 567).

3. Normative Constraint For Gilbert the plural will is created by individuals pooling their personal wills. Doing so does *not* mean that they are then left with no individual will, for our wills are not finite substances, but it does mean that (within the scope of the area of concern of the *we* in question) they are *rationally committed* to whatever the plural will is committed to; by voluntarily taking part in the plural will they become joined to it. The joint nature of this plural will means that it is under the control of the participants *together*; no one participant is individually in a position to unilaterally reject its demands by a simple personal change of mind or declaration. For example, Gilbert says of a couple walking together that, if one abandons the walk, the other might well rebuke them by saying “You can’t turn round, *we* said *we* would walk to the top!” The normative authority of the plural will arises *directly* from the nature of the joint commitment required for the construction of a plural subject – the plural will, formed collectively, can only be changed or rescinded collectively (see Gilbert 2008, 504). This gives rise to the following two criteria. Firstly, the obligation criteria which says that each participant has a pro tanto obligation to promote fulfilment of that which is intended by the plural will. Secondly, the permission criteria which says that participants understand that they are not (ordinarily) in a position to unilaterally “by a simple change of mind” remove the constraints imposed on them by the obligation criteria⁴. For Gilbert this normative constraint is not some additional feature of morality, peer pressure or politeness, rather it arises *directly* from the collective will – it is akin to the normative constraint that as individuals we face with our own intentions; e.g. if I intend to give this paper today but I actually stay in bed, then I have done something normatively wrong – likewise if *we* intended to together give a paper today but *we* stay in bed then we have done a similar wrong. The key difference between the plural will and one’s individual will is that we are in a position to unilaterally change our individual wills however each individual is not in a position to unilaterally change the plural will.

Westlund does not give a direct argument against the general possibility of such normative constraint⁵. However it is, she believes, contrary to her account of CRRs. In contrast to a plural will that each participant becomes joined to, Westlund’s (not fully spelled out) view is that the creation and maintenance of a plural will involves an ongoing ‘dance of union, separation and reunion’, that is, it involves individuals being engaged in an “evolving framework” of plural stances to which they must “*continuously reaffirm*”. Westlund does not discuss Gilbert’s scenarios of dissent, such

³ The example with a slightly different presentation is Westlund’s (2009, 1)

⁴ Gilbert (2000, 17) Though note, Gilbert allows that background conditions may allow an individual to control the plural will, e.g. “If one partner is discovered to have engaged in sexual activity with a third party, the offended partner may aver, ‘We’re through!’ and the other may not question that point” because of “... the existence of an established [background] condition” (2006, 110).

⁵ Merely suggesting that Gilbert’s claim is “controversial” and pointing to others (Bratman 1999, McMahon, 2005) for a general argument (Westlund 2009, 14, footnote 26).

as the two walkers above – if she did she would have to reject the picture Gilbert paints for her view implies that with regard to collective plans each participant must have “...some discretion over her own follow-through” (Westlund 2009, 14). That is not to say that Westlund wants to claim that the plural will does not have *any* kind of sway over each individual. If this were the case then it would be hard to see how it could play the role of making the reasoning of couples in CRRs fundamentally different from that of the game show bargainers. However, Westlund explains such sway by claiming that it exists because there is a *further* commitment to a robust form of mutual accountability.

We might think that Westlund needs to differ from Gilbert in this way in order to solve the problem of reciprocity. The concept of plural subjects is supposed to solve the problem of reciprocity because it allows the possibility of forming unions without losing each partner’s individual autonomy. Further, for Westlund a key part of this is that the joint deliberative agency, by which the couple form their collective, necessarily involves an ongoing sensitivity of each to the will of the other. This, it might be thought, rules out each party being directly constrained by the collective will, for their being so constrained might be seen as a block on each having such sensitivity. So for Westlund “... any balance that is struck between the parties must be regarded as defeasible in the face of further reflection and experience, such that a shared practical perspective is always a work in progress” (Westlund 2008, 568). I will refer to this in what follows as the *requirement of flexibility*.

4. The Requirement of Flexibility

There are two reasons to reject Westlund’s rejection of direct normative constraint; firstly because of its importance to Gilbert’s overall project, and secondly, because direct normative constraint *does* seem to be part of the actual phenomenology of being in a CRR. On the first point, a full defence would require an overall defence of Gilbert’s project that is beyond the scope of this paper but it is worth noting that the ability to explain the normative authority of plural subject attitudes as arising *directly* from the nature of joint commitment is what makes Gilbert’s account distinctly collectivist and thus able to explain the phenomenology that (she claims) individualistic accounts cannot.

On my second claim; it seems to me disputable that we experience CRRs as necessarily having the flexibility that Westlund describes. It is true that when a relationship is healthy each party tries to take the needs of the other into account when engaged in joint deliberation. In the holiday example we would say that if the couple, Mary and Clare, are in a well functioning CRR then each party should try to accommodate the desires of the other and to compromise when these do not fit with their own. Even once they have come to their collective decision, i.e. once they have constructed their collective will, then *if they truly care for each other* each will be open to the possibility of further deliberation over its content. This doesn’t mean that the collective will itself must be *automatically* reflective of any change in each individuals perspective. If Mary comes to realise that she hates hot nights, then this *does* mean that (given Clare’s love for her) Clare ought to be open to re-engaging in joint deliberation and to their jointly changing their mind, but it *does not* imply that the collective will must automatically cease to hold sway. In parallel with Gilbert’s walkers I suggest that if on the way to the airport Mary turned round and stated walking back home, Clare would be justified to say: “You can’t go home, we said we were going to Majorca!”

Further to Westlund looking for flexibility in the wrong place in healthy CRRs, she also idealises love by ignoring the fact that *unhealthy CRRs still count as CRRs*⁶. Feelings of being a *we* do not

6 A similar point (though not directed at Westlund) is noted by Gilbert, 1996.

necessarily disappear with the failure of each party to treat the other with reciprocity. Once we allow the phenomenology of badly functioning CRRs into our picture then the fact that there is no necessary flexibility at the level of the collective will becomes all the more apparent. Given that Westlund seems to believe that being bound to the collective will requires an additional personal commitment she would have to say that the shared perspective merely dissolves in the face of one, or both, parties ceasing to continuously reaffirm the framework which takes the other into account; but this is misguided as even in bad relationships we can feel the constraining force of the collective will.

5. **Rebelling and the Variability of Normative Pain**

While Gilbert gets closer to the phenomenology of love than Westlund, I will now argue there is *something* overly restrictive in her set up, just that it is not what Westlund thinks it is. My claim is that there is variance in the levels of what we might call “normative pain” when rebelling against the plural will of different CRRs. Just as we can understand physical pain as a negative feeling, of variable strength, experience of which indicates bodily damage, normative pain is negative feeling, of variable in strength, experience of which indicates normative transgression. This is, I believe, evident in the cases described below. Unfortunately, for showing this to be the case, the direct normative commitment of the will that Gilbert describes is not the only thing that provides reasons for the individual to feel tied to the *we* in cases of love – there are of course also moral, romantic, practical and further reasons in play. My claim is not that these other reasons do not matter – only that direct normative constraint of the will plays an important part in the mix. I’ll proceed by contrasting three examples to illustrate the variability of normative pain:

Early days: Imagine a newly formed couple, Bill and Ted. They have been together for a few months, seeing each other once or twice a week. They see themselves as very much in love, they are full of the strong feelings of lust and desire, but live in separate houses, have separate groups of friends and different hobbies.

Long term and going strong: Now contrast this with our couple from the example above, Clare and Mary. Let’s say that they have been together for 25 years, bought a house together, adopted and raised a child together, share friends and have the same hobbies. They feel still very much in love.

Coming apart: Patrick and Madeleine have been in a relationship longest of all, 30 years. They do live together, and pay bills jointly and have raised children. But they also do an increasing amount of things apart, have separate friendship groups and enjoy different social activities. They don’t really feel much romance toward each other and often find themselves attracted to other people.

Now imagine that each of these couples is on their way to the airport and one of each of them turns round and starts walking back home. Each rebel will be open to a rebuke of Gilbert’s form: “You can’t go home, *we* said *we* were going to Majorca!” However, I suggest that the act of rebelling, will feel different (in a relevant way) to each; When Ted turns round to Bill and announces that he is not going, the wrong that he is committing by violating the collective will feel less normatively painful than that which will be felt by Mary in doing the same. Similarly, the wrong of Madeleine’s doing so will feel worse than Ted’s, though not as normatively painful as Mary’s deviation.

I am not claiming that any of the couples face *no* obligation towards their collective wills, rather that the degree of normative pain of breaking these obligations will vary. It is the existence of this phenomenon and the way in which it mimics a certain aspect of Westlund’s description of flexibility that accounts for a certain level of intuitive plausibility to Westlund’s rejection of

Gilbert's account of direct normative constraint. Of course this could all be a function of variance in other factors; the different level of moral, romantic, practical obligations felt by each party. However, that the rebuke may feel more normatively painful to Madeleine (who is not getting that much out of her relationship any more) than to Ted (who is infatuated) suggests to me that the difference lies in the nature of the *we* rather than the personal feelings of each party. Further, given that the rebuke makes *direct* reference to the will of the *we*, I think we should look for variance towards this relation to explain this phenomenon.

So, how can we make sense of the variability between the couples? In what follows I give a rough sketch of the type of account I think could do this⁷. Such an replaces Gilbert's discrete, all-or-nothing, voluntary pooling of the wills with a continuous notion of wills becoming entangled through levels of necessary identification with the plural subject. The variance in the three holiday cases will thus be explained by variance in the amount to which the individuals cannot help but identify with the collective, where this variance will in turn be explained by how much of their own lives have been lived, and thus only fully *make sense*, within the scope of the collective agency.

The background to this claim is an understanding of what it is about commitments of the will in general that bind us. What is wrong, one might ask, with agentive anarchism, i.e. doing whatever one feels the compulsion to do at any time and not feeling committed by what one has willed? According to Michael Bratman our seeing ourselves as being constrained by the commitments' of our wills is necessary for us to be able to understand ourselves as agents who can govern our own lives⁸. This is because the commitments of our wills provide the scaffolding necessary to construct a place "where-*I*-stand stand" out of the multiple of elements of our psychological stew. I propose that we can extend this claim to explain the normative power of commitments of the collective wills for couples in CRRs; i.e. that each lover must see themselves as constrained by the will of the *we* that they are part of in order to see that *we* as a unified agent that can act⁹.

The case is not completely symmetrical between individuals and collectives. For individuals one always needs to have a where-*I*-stand as if we can't see ourselves as individual agents there is nowhere left to retreat. It follows that there is no variability in the normative pain of breaking with our own wills¹⁰. For a collective we *do* always need to have a where-*we*-stand for the possibility of collective agency. However, there is at least some possibility of retreating back to our own individual agency. What stops such a retreat being *too* easy is that part of our understanding of ourselves will be bound up in being able to see the collective as acting. Insofar as we have already lived part of our lives through the collective we can only continue to understand our contributory action as the kind of thing we set it out to be if we are able to see the collective as an agent, and because this requires its commitments of will to constrain, we must see them as doing so¹¹. This difference between the foundation of the power of the individual will and the collective will is, I believe, what explains the variability in the three holidaying couples. The variability comes from the fact that though we will always have some *part* of ourselves bound up with the *we*, the amount of this entanglement can vary (with length of time of collective etc.).

⁷ The argument here is condensed because of limitations of space, I proving an extended account in Kisolo-Ssonko (*forthcoming*).

⁸ Bratman (1999, 2007).

⁹ Note that this is not a use that Bratman has need to employ for he rejects the idea of direct normative constraint by a collective will.

¹⁰ Or at least if there is variance it is not of this sort.

¹¹ Post-hoc reconceptualising our contributions is logically possible but seems to be a kind of inauthenticity for that is how, at the moment of our actions, we set them out to be.

Acknowledging variability in the experience of normative pain requires a re-understanding of Gilbert by moving away from voluntary wills being pooled to wills becoming entangled over time perhaps in a sub-voluntary way, but I contend that this fits better with our phenomenological experience of being in love. Imagining one's life without one's significant other, imagining the non-existence of this *we*, is thus a conceptual issue rather than just an epistemic one – and this is how it grounds normative constraint.

REFERENCES

- Bratman, M. (1999), "Shared Cooperative Activity", in his *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge University Press, Cambridge, 1999, pp. 93-108;
- Bratman, M. (2007), *Structures of Agency: Essays*, Oxford University Press, Oxford, 2007;
- Gilbert, M. (1992), *On Social facts*, Princeton University Press, Princeton, 1992;
- Gilbert, M. (1996), "Fusion: Sketch of a contractual Model", in her *Living Together: Rationality, Sociality & Obligation*, pp. 215-228, Roman & Littlefield, Maryland. Originally published in *Perspectives on the Family*, eds. Moffat R. et al, pp. 65-78, Edwin Mellen Press, Lewiston;
- Gilbert, M. (2000), *Sociality and Responsibility*, Roman & Littlefield, Maryland, 2000;
- Gilbert, M. (2006), *A Theory of Political Obligation*, Oxford University Press, Oxford, 2006;
- Gilbert, M. (2008), "Two approaches to shared intention: An Essay in the Philosophy of Social Phenomena", *Analyse & Kritik*, Vol. 30, 2008;
- Kisolo-Ssonko, J. (*forthcoming*), "How Where I Stand Constrains Where we Stand", *Proceedings of Collective Intentionality VII*, Studies in the Philosophy of Sociality, Springer, Dordrecht;
- McMahon, (2005), "Shared Agency and Rational Cooperation", *Noûs*, Vol. 39, no. 2, pp. 284-308, 308, 2005;
- Nozick, R. (1995), "Erotic Love", in *Philosophical Perspectives on Sex & Love*, ed. Stewart, R. Oxford University Press, Oxford, 1995;
- Roth, A. (2004), "Shared Agency and Contralateral Commitments", *Philosophical Review*, Vol. 113, pp. 359-410, 2004;
- Searle, J. (1990), "Collective Intentions and Actions", *Intentions in Communication*, ed. Philip Cohen et al, M.I.T Press, Cambridge, 1990;
- Tuomela, R. & Miller K. (1988), "We-Intentions", *Philosophical Studies*, vol. 53, p.367-389, 1988;
- Velleman, D. (1997), "How to Share an Intention", *Philosophy and Phenomenological Research*, vol. 57, no. 1, pp. 29-50, 1997;
- Westlund, A. (2008), "The Reunion of Marriage", *The Monist*, Vol. 91, No. 3 & 4, 2008;
- Westlund, A. (2006), "Love and the Sharing of Ends", *Twenty-First Century Papers: On-line Working Papers from the Center for 21st Century Studies*, University of Wisconsin, Milwaukee, 2006;
- Westlund, A. (2009), "Deciding Together", *Philosophers' Imprint*, Vol. 9, No. 10, 2009.

NEIL W. WILLIAMS

University of Sheffield

neilwilliamwilliams@gmail.com

AGAINST ATOMIC INDIVIDUALISM IN PLURAL SUBJECT THEORY

abstract

*Within much contemporary social ontology there is a particular methodology at work. This methodology takes as a starting point two or more asocial or atomic individuals. These individuals are taken to be perfectly functional agents, though outside of all social relations. Following this, combinations of these individuals are considered, to deduce what constitutes a social group. Here I will argue that theories which rely on this methodology are always circular, so long as they purport to describe the formation of all social groups, as they must always presuppose a pre-existing collectivity. Such methodology also produces various distortions in our theories, such as voluntarism. I focus on the workings of Plural Subject Theory as laid out by Margaret Gilbert in *On Social Facts* (1989). I show that the formation of a plural subject always requires communication, and that communication always requires a pre-existing collectivity. I examine the elements within Plural Subject Theory which protect Gilbert from these accusations of circularity, and argue against them. I finalise by suggesting that what Plural Subject Theory, and social ontology in general, requires as a theoretical starting point is not atomic individuals and their combinations, but rather combinations of already socialised or embedded individuals.*

keywords

Atomism, plural subject theory, voluntarism, social ontology, group formation

“A brief hypothetical account of the transition from the state of nature to a social state ... First nature (or God) produced a rational or reasoning being. Such being moves by virtue of acts of will ... This being was endowed with ideas of its own desires such that ‘I want a drink’ was understood to provide reason to move towards a drink. Now more than one being of this kind was produced. What ideas could these beings be endowed with that might enable them to live together harmoniously?”
(Gilbert 1989, 414).

1. Introductions and Distortions

Theorisation necessitates abstraction: to gain traction on the messiness and complexities of reality we must abstract away from it. Not all abstraction is legitimate however, and one must be assured in an investigation that the chosen abstractions do not distort the aspect of reality under examination. This paper argues that such a distortion is at work within the methodology most common in contemporary analytic philosophy of society. The abstraction in question is neatly demonstrated in the epigraph: though its story is “crude” it is none the less taken to “make sense” (OSF, 415). Such methodology relies on the assumption that one can legitimately abstract away from society a *primordial atomic individual*, and then see what combinations of these individuals create a social group.

“Methodological Individualism” was first suggested by Weber as a scientific and “un-ideological” basis of a science of society (Weber 1922/1978, 13-16). It was simply the assertion that society should be understood via the intentions, actions and beliefs of individuals. In actual fact, the methodology at work within much contemporary philosophy of society is more akin to methodological atomism understood in a Hobbesian sense, which sees human beings as capable of coming into full maturity outside of any societal interaction (Hobbes 1651/2003, 102). Whether individualism or atomism the important methodological assumptions I am questioning are these: (i) that we can feasibly consider an asocial human agent as our primary unit of investigation, and (ii) that we can consider combinations of these individuals to deduce what is intuitively required to produce a “social group”. With these methodological commitments, the main question of societal philosophy becomes: “What makes a collectivity out of a sum of living beings?” (Gilbert 1989, 2).

I take it that, by definition, the majority of “summative accounts” take this methodology to be correct. Summative accounts argue that all talk of groups can be reduced to talk of individuals, and that groups don’t *really* exist as anything other than a combination of individuals¹. Methodological atomism makes the summative theory of groups very plausible – a methodology that places such weight on the atomic individual will be more likely to produce an individualistic ontology of society as an output. In this paper however, I focus exclusively on the “Plural Subject Theory” (PST)

¹ Summative accounts include Quinton (1976), Keeley (1981) and Bratman (1999). See Gilbert (1989, 257-288) for discussion on the nature of summative accounts in relation to a plural subject account.

forwarded by Margaret Gilbert. PST is a rarity: a *non-summative* account of social groups achieved through investigative atomism of the kind sketched above. A plural subject is a kind of collective intentionality, a subject made by a “unity”, a “pool or sum of wills dedicated, as one, to a certain ‘cause’” (Gilbert 1989, 409). Plural subject-hood is not reducible to the singular intentionality of the persons so connected². I focus on Gilbert’s account precisely for this reason, for I take it that if I can show that Gilbert’s non-individualistic account shows distortions due to her atomistic methodology, then summative accounts certainly will. In particular, focus will be on Gilbert’s *On Social Facts* (1989), which stands as the formation and bedrock of PST³.

Finally, I must indicate the nature of the distortion such methodology creates. Here I focus mainly on a distortion which can be called *voluntarism* within Gilbert’s theory. As shall be shown, the methodology at work leads Gilbert to investigate almost exclusively small voluntaristic groups of two or three people, who perceive themselves to be a unit. Examples include mushroom pickers (Gilbert 1989: 36-41), social travellers (Gilbert 1989, 161-164) and – importantly – conversers (Gilbert 1989, 433). Gilbert takes these small groups as “paradigmatic” of all social groups (Gilbert 1990), and as the origins and beginnings of all larger groups (Gilbert 1989, 235)⁴. The formation of these groups are the formation of plural subjects (See §2). Thus the conclusions drawn from investigating these small, voluntaristic groups can be extended to *all* social groups (Gilbert 1989, 149). Gilbert accepts this voluntaristic nature as a necessary part of PST: “if there is a plural subject, *then* each set of persons has *volunteered* his will for a sum of wills” (Gilbert 1989, 413, emphasis mine). See ing as all social groups are plural subjects (Gilbert 1989, 204, 233), this means all social groups are voluntaristic. This leads to odd theoretical conclusions such that we can always leave a group (Gilbert 1989, 426), and that those in power are always somehow representative of the people (Gilbert 1989, 206). Gilbert admits and highlights this conception of society with frequent comparisons to Rousseau’s social contract theory (Gilbert 1989, 198, 206, 415-6; 2008: Chapters 4, 5 and 10). At best such voluntarism creates a theory that is bias towards the ideal of western, democratic societies, and at worst it avoids, trivialises or removes from investigation families, organisations and societies that are repressive, coercive or dictatorial.

At a theoretical level, such distortions may seem harmless, but Gilbert’s primary aim is to develop a collectivity concept that can be applied in the social sciences (Gilbert 1989, 2-3), and which can be extended to (at least) sociology, anthropology, political philosophy and psychology (Gilbert 1989, 436-41). Indeed, such distortion is carried over into Gilbert’s more recent work developing a PST account of political obligation (2008). The PST account argues that membership to a group is *always* at some level intentional (2008: 168), that *any form* of government can be backed by joint commitment (2008, 180) and that a population has an obligation to comply with political institutions *whatever they are* (2008, 256). As such, these distortions should be taken seriously.

Primarily, I focus on conversation as an example of the formation of the plural subject. I combine contemporary linguistics with Gilbert’s theory, to show there is a circularity inherent in her notion of plural subject-hood (§2). I argue against Gilbert’s objections that language is an asocial phenomenon (§3), and suggest in conclusion that the best way to break the circularity in question is to concede that society is not a plural subject (§4). Throughout, I attack the notion of the atomic individual, and suggest that instead, PST should take *the already socialised individual* as its primary investigative unit.

2 Plural subjects will be better defined in section 2. Gilbert forms PST against the assumptions of “singularism”, which sees individual intentionality and agency as the only kind of intentionality and agency there can be (Gilbert 1989, 12).

3 Readers will be referred to more recent examples of Gilbert’s work when necessary.

4 Gilbert maintains this dedication to the importance of small groups in more recent work (Gilbert 1990; 2000, 14-36; 2008, Chapter 6). Due to the methodology under question, such small units of investigation are common in contemporary philosophy of society (see Bratman 1999, 93-108).

2. **Conversations
and Actions**

When discussing the *formation* of plural subjects, Gilbert investigates how “normal human beings with whatever capacities these have” combine to form plural subjects, the proper referent of “we” (Gilbert 1989, 175). This is methodological individualism in action: an investigation into how human agents outside of a plural subject combine to form one. I focus on the example of *conversation* for two reasons. Firstly, it is an oft used example (Gilbert 1989, 170, 200, 410, 433; Gilbert 2008, 101, 118, 173) of the small and “ephemeral” (Gilbert 1989, 215) nature of the social groups Gilbert investigates. Secondly, Gilbert makes it a requirement that agents must have “made certain things clear” (Gilbert 1989, 162), prior to the formation of the plural subject. This indicates she needs some notion of communication prior to the plural subject, or for conversation to be the primary plural subject. This being the case, problems with conversational plural subject-hood will resonate through Gilbert’s theory.

There are four stages to the formation of the plural subject, each of which must be met in turn for the plural subject to be formed:

1. *Quasi-readiness*: Each agent must be independently prepared or “set up” to engage in a joint action A with another agent. This is “a state or disposition of the will, a kind of readiness to act when appropriate conditions obtain” (Gilbert 1989, 186).
2. *Common Knowledge*: The quasi-readiness of each participant must be communicated, must be what Gilbert terms “open*” (asserting that humans have the concept of “openness*” by default (Gilbert 1989, 191)). “It is *common knowledge* among A, B and C that p, if and only if (by definition) the fact that p is open* to A, B and C, and (2) A, B and C have noticed this” (Gilbert 1989, 195).
3. *Joint Readiness*: An addition of (1) and (2): “a set of persons are jointly ready to share in action A in circumstances C if and only if it is common knowledge among them that they have mutually expressed their quasi-readiness so to share” (Gilbert 1989, 198).
4. *Pooling Wills*: This is the creation of a plural subject: “[O]ne who expresses quasi-readiness to do A in C in effect *volunteers his will* for a pool of wills to be set up so that in certain circumstances, that pool will be dedicated to a certain end” (Gilbert 1989, 198)⁵.

With these stages in mind, I turn to a contemporary theory of communication, Relevance Theory (RT), forwarded by Sperber and Wilson (1986/1995). In RT communication occurs by a speaker making information accessible in the audience’s “cognitive environment”. The cognitive environment is a combination of the agent’s *cognitive abilities* (the biological, perceptual, linguistic and conceptual abilities that agent has) with the *physical/sensory environment* that agent is located in. The cognitive environment is the set of facts and assumptions accessible or “manifest” to the individual. Communication occurs in the overlap between two or more cognitive environments. It works by making certain things *mutually manifest*, which is a useful way of thinking of Gilbert’s “openness*” requirement (See Sperber and Wilson 1986/1995, 36-45).

The combination of the two theories is illuminating. To begin a conversation, agents must be receptive and ready to engage one (1). This quasi-readiness must be communicated in order to attain (2) and (3). RT, as an *ostensive* theory of communication, sees any act of communication as carrying with it a necessary recognition of the participant’s *intention* to communicate. Communication is not accidental; by beginning to communicate we make our readiness to communicate mutually manifest. Something mutually manifest is common knowledge (2)⁶. The overlapping of cognitive environments is an excellent account of “joint readiness” (3). This creation of a *mutual cognitive environment* perhaps

⁵ More recent incarnations of PST keep these distinct stages (Gilbert 2008, 96).

⁶ Common knowledge is a complex topic, but I take it this RT account of common knowledge can be successfully applied to most, eg Lewis (1969, 52-57).

explains Gilbert's contention that we need to be considered an "us" in a sense *before* we embark on a full "pooling of wills" and a shared goal ("I suggest that A and B must need to be in a position appropriately to think of themselves as 'us' in order for A to intelligibly to set out to tell B something" (Gilbert 1989, 215)). The two agents can then embark on the joint goal: to converse (4) (Gilbert 1989, 170). Conversing is the manipulation of this shared cognitive environment, the making mutually manifest certain facts, assumptions, problems, and so on.

This account transfers to non-conversational joint action. Take moving a sofa together. The sofa must be manifest in my individual cognitive environment as something that I am ready to move with you (1). I must then make this intention manifest to you, and vice versa (2). The fact that this is manifest to both of us means it is *mutually manifest* – we have created an overlap of cognitive environments (3) and can proceed to the joint action of moving the sofa together (4).

Considering contemporary theories of communication sheds much light on Gilbert's account. RT not only fits with PST, but it elaborates the notions of communication that PST requires for its account of plural subject formation to be plausible. However, RT works on the basis that communication presupposes a commonality in *cognitive abilities* which allows participant's cognitive environments to be similar enough for overlap to occur. Such cognitive abilities are usually considered linguistic or conceptual and require a society to bestow such commonality. For PST a society is a plural subject, and so for the formation of a plural subject to presuppose society would entail circularity. To avoid this circularity, PST must show that the cognitive abilities required for plural subject formation are *asocial*.

For a plausible assessment of Gilbert's methodology, one must first distinguish what cognitive abilities an asocial human agent would possess, and what such an agent can only gain through societal interaction. Searle (1995, 127-47) distinguishes between *Deep* and *Local* "Background" in this regard, using the terminology to separate the capacities and abilities we have *naturally* (Deep) and those which we have *culturally* (Local). Though Searle focuses on capacities, abilities and predispositions, I focus solely on which *concepts* (or "cognitive abilities") can be considered to be natural. It is the placement of certain concepts into Deep or Local Background which is crucial to this discussion⁷.

Deciding which concepts are natural, and which cultural, is a matter of discussion and discovery⁸. But, to prevent us from simply placing concepts where our theories need them to be, we can present certain feasible criteria. I suggest that the primary candidates for "natural" concepts would be (i) something which can be said to arise from the ordinary use of our biological and/or sensorimotor abilities. Lakoff and Johnston (1980, 56-60) call these *naturally emergent* concepts, and suggest spatial and temporal concepts as paradigm cases (UP, DOWN, FORWARD etc.). To this I would add (ii) concepts which could arise through the perception of and interaction with a non social environment (LIGHT and DARK, perhaps DAY and NIGHT or natural kinds such as MOUNTAIN). *Deep Background* will contain the concepts that are indicated by these criteria. *Local Background* will contain the concepts we can only gain by virtue of being in any society, or a particular society.

The majority of Gilbert's dyadic examples presuppose conversation, and thus language. I will assume, for the moment, that a language presumes a society (See §3). What possible plural subject-hood

7 I borrow Searle's terminology but not his notion. There are deep flaws in Searle's conception which I do not wish to inherit. Suffice to say, I am committed to at least a significant part of the Background being conceptual and linguistic, which Searle's non-representational Background does not allow for.

8 I am drawing a theoretical line here, not an actual one. It sees *ms* "Local" Background will always shape and effect "Deep" Background. Recent empirical data shows that different cultures have radically differing conceptions of seemingly asocial concepts such as SPACE (Nisbett and Miyamoto 2005). Mallon *et al* (2009) recently used such evidence to argue against contemporary philosophy's reliance on *intuition*, an argument which would find much purchase in Gilbert's theory.

is possible without language? That is, what *shared cognitive environments* are possible? If I teleport a tribesman from a tribe as yet unaccountably and blissfully ignorant of Western culture, could I get him to enter into a plural subject with the simple joint intention to move a sofa? See ing as we share only Deep Background, how could I communicate my intent, and so get from quasi-readiness (1) to a plural subject (4)? Perhaps by some fluke of gesture, or by virtue of recognition of shared Deep Background, I could make my intention mutually manifest, and so entreat him to “join forces” (Gilbert 1989, 163). Beyond this most rudimentary physical joint action, what actions could we participate in without significant education on his part or mine?

The majority of Gilbert’s dyadic examples presuppose that agents have *previously been in a society*, which is even more of a problem. The teleported tribesman and I shared one salient piece of Local Background: that we both *had* a Local Background. We were both born and raised within a society. If I instead teleport a Wildman, who has never before seen a human soul, how would I go about entering into a plural subject with *him*? Would he know what it is to *cooperate* or *communicate*? Had I accidentally teleported him into a tiger’s pen, and was forced to communicate the salient fact that there was a tiger sneaking up on him from behind, could I? Communication relies on the recognition of *intention*. Would the Wildman have even this concept? Would he know that my desperate gestures and shouts were the act of a human *agent* intended to *convey* something? Judging from the prior criteria, the answer would be no. *Cooperation, communication, recognition of intention*, each necessary to even begin *thinking* about joint intention (1), have no reason to emerge spontaneously in the Wildman’s Deep Background. We would have to (somehow) embark on a great deal of education before this poor man and I could even begin to make things mutually manifest to one another. But what would this education be if not *socialisation*?

This then, is my central point: in order for PST to work it must presuppose that participants are (or have been) inhabitants of a larger, societal, plural subject. In order for PST to describe anything but the most rudimentary joint actions, it must presuppose participants live in the *same*, or relevantly similar societies. Simply put: the *more* cognitive overlap we share, the *more* joint actions we can engage in, and the *easier* it is to form a plural subject. If there is *no* cognitive overlap, we *cannot* engage in the formation of a plural subject.

Gilbert is quite comfortable with the notion that conversation (Gilbert 1989, 154), or at least communication broadly construed (Gilbert 1989, 215), is prior to and required for group formation. This does not lead to a “paradox” for Gilbert, who sees the possibility of a linguistic social isolate quite plausible (Gilbert 1989, 217, See §3). However, what investigating PST through the lens of contemporary communicative theory has revealed is that all communication requires an overlap of “cognitive environment” which requires salient commonalities in the Background. PST cannot account for shared Background without presupposing that the inhabitants already inhabit a social collective of the type PST was meant to explain. What this circularity reveals is that for PST to work it needs, and in some sense already takes, the notion of an *already socialised individual* as its primary methodological unit, rather than the atomic individual. I now turn to objections to the claim that communication is necessarily social, before sketching a non-circular PST in the final section.

3. **Objections and Abstractions** The plural subject theorist could answer my accusation of circularity with two objections. Firstly, one might say that a participant does not need to have *inhabited* a society (or other plural subject) in order to form a plural subject. A participant needs only the *concept* of a plural subject. On this claim the concept of a plural subject is hard-wired into us, part of our Deep Background. Gilbert suggests that the benevolent creator of the epigraph would need to bestow “an innate concept of a plural

subject” (Gilbert 1989, 167), or a certain kind of “mutual recognition” (Gilbert 1989, 217) to move from an asocial state to a social one. The role this innate concept of plural subject-hood plays in group formation is implicit through most of Gilbert’s work, but occasionally emerges explicitly:

I do, of course, posit a mechanism for the construction, so to speak, of social groups. And this mechanism can only work if everyone has a grasp of a subtle conceptual scheme, the conceptual scheme of plural subjects. Given that all have this concept, then the basic means for bringing plural subject-hood into being is at our disposal (Gilbert 1989, 416).

Gilbert needs to assert the nativism of a plural subject concept because it is the only way her account of group formation can work without circularity. There is no argumentation for this position, only assertion. I assume that we cannot simply assert innate concepts to save our philosophical system unless we have a reasonable causal story for how they would emerge naturally. Such an assertion is certainly not a defence against a claim that it is the methodology which creates the theoretical gaps which these nativist assertions are made to plug.

“Plural subjecthood”, as a concept, does not meet the criteria I sketched above for naturally emergent concepts. It is not a result of our sensorimotor abilities, and an asocial agent has no need or ability to produce it in engagement with an (asocial) environment. For Gilbert’s account to work the concept must just be a “given”. The burden is on Gilbert to explicitly argue why this is the case. Moreover, assume for a moment that Gilbert is correct. Asserting that human beings have an innate concept of collectivity is tantamount to asserting that human beings, *qua* human beings, are social creatures. Human beings have evolved within, or are otherwise specifically adapted for, social groups. Why then should philosophy take as the methodological unit of investigation an *atomic* and *asocial* human? Is it not bound to create distortions in a theory if the primary methodological unit is an animal outside of the habitat it is adapted to, a human normally considered at best an oddity and at worst pathological? The second, more substantial objection that the plural subject theorist could offer is the assertion that language is not necessarily social. If language is not necessarily social, then the fact that communication is necessary for plural subject formation does not have to be circular. Gilbert clearly saw how crucial it was for PST to develop an asocial theory of language, dedicating a whole chapter to the task (Gilbert 1989, 58-146). This chapter is split between attacking Wittgensteinian arguments against private language, and asserting a PST account of group language formation. Here I pick out only one strand of the argument: the congenital Crusoe intuition. The congenital Crusoe figure is a social isolate with full grasp of a private language. For Gilbert’s account of language to function this figure must be plausible. Moreover, this fully functioning but socially isolated individual is precisely the atomistic individual which I am arguing against. If this figure is plausible, then the methodology in question is unproblematic. The Crusoe figure, perhaps because it is presupposed in her methodology, is a conceptual certainty for Gilbert:

the claim ... that it is logically impossible to have a language if one has not participated in group life ... [is] counterintuitive on the face of it. It appears we can at least conceive of a congenital Crusoe – a being socially isolated throughout his life – who was initially endowed with, or invented, a language of his own (Gilbert 1989, 59).

Gilbert assures us this is pre-theoretical or “untutored” intuition. It is “intuitive” for her that this figure can use language (Gilbert 1989, 73), and often if an opposing argument reaches the “counter-intuitive” “anti-Crusoe conclusion” then this “counts against it” (Gilbert 1989, 89-90), and one should become “immediately suspicious” (Gilbert 1989, 95) of such a theory. This intuition works as a motivating feature also: on developing what she finds to be an “intuitive” or “natural” model of language (Gilbert 1989, 93-4) – essentially Platonic – “one datum” in its favour is its avoidance of the anti-Crusoe conclusion (Gilbert 1989, 95).

The first question is to ask is whether the Crusoe figure is in anyway *actually* possible. Recent empirical evidence suggests that it is not. Evidence indicates that children who do not learn language during the “critical period” of development (around 3-13) do *not* develop full linguistic ability (Pinker 1994: 38-9). Without learning a language within this period, we lose the ability to learn *any* language completely, even when subsequently placed in a linguistic environment (Pinker 1994, 290-2). Pinker cites the case of “Genie”, a so called “wolf-child” who was discovered at the age of 13. Genie was without language and severely impaired in her ability to learn one (Pinker 1994, 291). Generative linguistics suggests that though we have a biological ability to develop language when in a linguistic environment, and within a certain key time frame, children that cannot interact with a language during this time *cannot* and *do not* develop one independently, spontaneously or by intellectual effort alone. Though Gilbert’s argument is primarily conceptual rather than empirical, I suggest we should at least be sensitive to empirical facts in our theorising, especially when developing concepts for use in the social sciences. Though the Crusoe figure may be a *conceptual possibility*, it is not a *nomological possibility*.

Indeed, the Crusoe figure may not be as *conceptually* possible as Gilbert thinks. When imagining the conceptual Crusoe figure, one imagines a socialised individual, *like us*, but without society. Of course it seems plausible that such a figure could develop language independently. Equally, if we picture the social isolate as *just like us* but outside of society, it seems very plausible to imagine that they could form groups with ease when other humans are introduced. Now try and consider someone with *only* Deep Background. Is it plausible that this person would spontaneously develop the tendency to attribute signs to objects? Can a human entrenched in society imagine the mental life of someone who has never before interacted with another human? It would be, doubtless, not at all *like us*. In order to even ascertain whether the Crusoe figure is conceivable, a great deal more fleshing out of this figure is required by Gilbert⁹.

Despite empirical invalidity, and conceptual inadequacy, Gilbert perhaps deserves an answer as to why she finds the Crusoe-figure so intuitive. I take it that Gilbert’s view, and that of Winch whom she argues against, fall down either side of the same hurdle. Winch gives us an individual incapable of intentional action outside of society. This is clearly undesirable, but in response Gilbert gives us an individual with a fully functioning vocabulary and conceptual apparatus, despite having never seen another soul. This is an equally radical view. I suggest that if we make a simple distinction, between *concept* and *language*, we gain a much more feasible picture. With this the appeal of the private language argument falls away. If one conflates *concept*, *language* and *intention* (Gilbert 1989, 68), the conceptual Crusoe argument becomes more plausible simply because it is unintuitive to deny the Crusoe figure *all* of these. But separate *concept* and *language* and one finds that we can have a plausible Crusoe figure possessing *concepts* and *intention* (with Deep Background) but not necessarily *language* (without Local Background). Gilbert requires an argument for a conflation which is simply assumed. The account of group language formation that Gilbert gives is very similar to the plural subject formation already looked at. It simply involves social isolates, each with their own private language, figuring out what words they shall use publically (Gilbert 1989, 132-142). Language formation, like group formation, is voluntaristic and intentional¹⁰. It *only* makes sense if we already assume the linguistic capacities of the social isolate. Gilbert then begs the question when she assumes that the congenital Crusoe “has a language”. Against a similar accusation Gilbert answers:

If we start from the standpoint of what is acceptable intuitively, however, there is no need of an argument to show that those who write off language-using isolates have no right to do so. In other words, this question may indeed

9 Considering Gilbert’s earlier paper on asocial language (1983), Sharrock and Anderson have a similar response: “Doubting whether someone can invent a language for themselves involves raising rather more deep and complex questions than those which can be met by the response ‘Well, I can imagine someone doing it!’” (1986, 554).

10 Chomsky, in particular, has staunchly denied the empiricist conception that language formation is an intellectual achievement, preferring the notion that it is akin to organ growth (See Pinker 1994, 297-332). Chomsky’s ‘poverty of the stimulus’ argument focuses on why it is implausible to consider language an intellectual creation (See Chomsky 1980, 310-24 and 2000, 3-19; Cook 1991; Pinker 1994, 262-297).

be begged, until it is shown that such statements are incoherent or otherwise inadmissible. More precisely, some flaw must be clearly shown in the picture [of congenital Crusoe] that I have sketched (Gilbert 1989, 99).

Such a statement amounts to little more than an assertion that her intuition is correct. I am now in a position to answer the challenge here laid out with not one but three “flaws” in the congenital Crusoe conception. Firstly, it goes against contemporary empirical evidence. Secondly, it is not conceptually motivating without more work. Thirdly, it rests without argument on a conflation of concept and language. Though it is plausible to assume that humans are *conceptual* and *intentional* creatures outside of society, there is no reason to think that they are *linguistic* or *communicative* creatures outside of society, which is what Gilbert needed to provide to avoid the circularity I identified. This places the burden of proof rather forcibly back onto the plural subject theorist.

This examination of the implausibility of the congenital Crusoe figure doubles as an argument against methodological atomism. Taking an empirically impossible and conceptually flawed unit as the basic building block of a theory is going to lead to the distortions I have illuminated, and (no doubt) others. Once again I suggest that what PST and other theories of social ontology require as the primary methodological unit is the *already socialised individual*. This amounts to accepting the circularity I have shown. Plural subjects require an already present collectivity for their formation. In the final section I offer one plausible way for PST to diffuse this circularity, and the distortions the methodology entails¹¹.

The primary methodological unit needed for PST to work must be the *already socialised individual*. As this is conceptually and nomologically plausible, I take it to be preferable to that which it replaces: the non-plausible atomic individual that much philosophy of society presupposes. This ultimately amounts to little more than the argument that we should take human beings as they have always existed – as entrenched in social reality – as our primary investigative unit in the social sciences¹². I take it such a position would be uncontentious if it wasn’t for the circularity which it apparently entails. So I now finalise by sketching a non-circular and non-distorted research possibility for PST. I suggest that the simplest way to break the circularity inherent in Gilbert’s PST is to deny that society is a plural subject. This is not to say that PST doesn’t pick out an interesting, important and useful notion for the social sciences. It is to deny that *all* the social sciences need is PST, or that *all* social groups are plural subjects (Gilbert 1989, 413; 2008, Chapter 8). PST describes exactly what it proclaims to describe: small, voluntaristic, united and self aware groups. Social groups of the sort united by goals or actions. But, as argued above, a) such groups presuppose collectivity, so they cannot be the sole collectivity concept, and b) understanding all of society in terms of PST leads to distortions such as voluntarism (1).

4. Options and Conclusions

¹¹ In order to clarify the argument here it will serve to compare it with the useful distinctions put forward by Pettit (1993). Pettit suggests that there are two separate issues that are often conflated in social ontology: the individualism/collectivism issue (1993, Chap 3), and the atomism/holism issue (1993, Chap 4). The collectivism/individualism question revolves around whether our intentional agency is impaired by social regularities (the collectivist position), or unimpaired (the individualist position). As I understand it my paper is neutral on this issue. I am concerned here about the issue of voluntarism only in as far as it pertains to the suggestion that we in some real sense *choose* to enter the society or family we find ourselves in. I am neutral here as to whether our ability to act autonomously within these social groups is impaired in any way. This being said, I think there are many issues that come to the fore when we consider our lack of choice in this regard, which may affect our autonomy. Pettit avoids these questions by narrowing his investigation to “autarchy”, or the minimal requirements for an agent having rational intentional states (1993, 120). It may surprise the reader to learn that the paper is also neutral on the atomism/holism issue, at least as it is described by Pettit. Pettit’s concern is whether interaction with others is required for full human rationality (the holist position), or whether an asocial creature can develop full rationality outside of human interaction (the atomist position) (1993; 1998). Though we both argue against atomism, we do so for different reasons. It may be plausible to suggest that a truly asocial individual is still fully capable of thinking, even if impaired in any number of other ways. I have suggested that an asocial individual would still have the resources to be an intentional individual (“I see no way of arguing that a Robinson Crusoe, even a Robinson Crusoe isolated from birth, could not think or follow rules” (Pettit 1993, 179)). The point at issue within this paper is whether or not such an asocial individual would have the cognitive resources to form, or participate within, a social group. The answer to that, I have argued, is no. In this sense I am a holist about group formation and participation, even if I am not (necessarily), one about human rationality. There is, no doubt, reason to believe that holism about one leads to holism about the other, but this is a discussion best left for another day.

¹² The position I have arrived at here is, as I take it, not dissimilar to the intuitions relating to embedded individuals which drive the phenomenological investigations of (for instance) Heidegger and Merleau-Ponty, as well as the work of the Pragmatist John Dewey.

So how *does* the formation of a society work, if it is not described by PST? I take it this is a question that philosophy of society does not have to answer: human beings and society evolved together, and the two are inextricable in a real sense. A human (somehow) born or raised outside of social interaction is – as poor Genie shows – severely linguistically, socially and emotionally disabled. The question should not be “how do societies form, given atomic individuals?” but rather “how do societies function and develop given the interactions of embedded individuals?”

So what *is* society, if it is not a plural subject? Well, this is hardly a question I am expected to answer here, but this exegesis of PST gives some idea of how to move forward. Society is *not* a “social contract” (Gilbert 1989, 198) in the way that Rousseau and PST formulate. Society is something we are born into, not something we choose. Society is better understood as the systems, structures, institutions and hierarchies in place. These affect those individuals born into the environment by shaping their subjective values, expectations, language and concepts: creating the *Local Background*. Already socialised individuals who share overlapping Backgrounds are capable of forming plural subjects. Plural subjects are the agents of group action within a society. Society is *not* a plural subject, but a different kind of collectivity which is the pre-condition of plural subjecthood. Society is not a plural subject but a system of overlapping and interacting plural subjecthoods. This is the merest suggestion, but shows how one might progress in incorporating PST into a less distorted investigation of social interactions.

In response to this picture, the plural subject theorist might offer two objections. The first is to claim that our society does show distinctive “plural subject” tendencies, where the population as a whole has collective goals or actions. A prime example often given is voting¹³. This could be construed as a more normative claim: “Okay”, the plural subject theorist may concede, “most *actual* societies are not plural subjects. But they *ought* to be plural subjects, on the basis of our common collectivity concepts, when they are functioning properly”¹⁴. This would then be a theory which distinguishes between political institutions at the theoretic level, and can only recognise and investigate more (idealised) democratic societies as social groups. This would mean that the vast majority of existent societies could not be investigated by the social sciences, which is not a desirable consequence¹⁵.

The second objection stems from Gilbert’s investigation of the “Kafkaesque” company. Such an organisation – full of employees unaware of one another and without collective thought – is simply *not* a “social group or human collectivity” on Gilbert’s account (Gilbert 1989, 231). The analogous society would be an incredibly dictatorial or repressive regime. If one reads Gilbert as saying that the alienating organisation and dictatorial regime are *not social groups* then the above criticisms apply. If one reads Gilbert as saying they are *not plural subjects* then this essentially agrees with my claim: many (most) societies and organisations are not plural subjects.

I conclude that methodological atomism leads to distortions in PST, which are overcome when the atomic individual is denied. Without this atomic individual the voluntaristic plural subject model of society is no longer plausible. A narrowed but undistorted PST is then free to describe the formation and activity of voluntaristic social groups embedded within society. A final note is required for the plural subject theorist who may concede that I have identified and overcome the circularity and distortion within PST, but at too high a cost. For the picture now is a truly *un-voluntaristic* account, an account that is in danger of losing the agency the original model granted the individual by definition. In response, I suggest that this is not a *theoretical* problem, but a *practical* one. Individual agency within a collectivity is not a given, but something which must be *strived* for, part of that “ancient problem of how to live” (Gilbert 1989, 436)¹⁶.

13 Considering the percentage of the population that actually votes decreases with each election, I believe the claim that society is functioning as a plural subject at these times is not without significant empirical problems.

14 Gilbert (2008, 180-181) offers a similar suggestion.

15 I leave open the possibility that our common collectivity concept is that of the plural subject, and that society is *not* a plural subject, but through misapplication or even manipulation we often mistakenly *think* of society in plural subject terms.

16 A shorter version of this paper was delivered at the *Collective Intentionality VIII* conference (Manchester, 2012), and I am grateful to the comments which it received. I am indebted to Dominic Gregory and especially Joseph Kisolo-Ssonko for their assistance with earlier drafts.

REFERENCES

- Bratman, M. (1999), "Shared Cooperative Activity" in *Faces of Intention: Selected Essays on Intention and Agency*, Cambridge University Press, Cambridge UK, 1999, pp. 93-108;
- Chomsky, N. (1980), "Discussion of Putnam's Comments" in Piatelli-Palmarini, M. (ed) *Language and Learning: the Debate between Jean Piaget and Noam Chomsky*, Harvard University Press, Cambridge MA, 1980, pp. 310-24;
- Chomsky, N. (2000), "New Horizons in the Study of Language" in *New Horizons in the Study of Language and Mind*, Cambridge University Press, Cambridge and New York, 2000, pp. 3-19;
- Cook, V.J. (1991), "The poverty-of-the-stimulus argument and multi-competence", *Second Language Research*, 7(2), 1991, pp. 103-17;
- Gilbert, M. (1983), "Has Language a Social Nature?", *Synthese*, 56, 1983, pp. 301-318;
- Gilbert, M. (1989), *On Social Facts*, Princeton University Press, Princeton, 1989;
- Gilbert, M. (1990), "Walking together: A Paradigmatic Social Phenomenon", *Midwest Studies in Philosophy*, 15, 1990, pp. 1-14;
- Gilbert, M. (2000), *Sociality and Responsibility: New Essays in Plural Subject Theory*, Rowman & Littlefield Publishers, Lanham MD and Oxford UK, 2000, pp. 14-36;
- Gilbert, M. (2008), *A Theory of Political Obligation: Membership, Commitment and the Bonds of Society*, Clarendon Press, Oxford, 2008;
- Hobbes, T. (1651), *On the Citizen*. Tuck, R. and Silverthorne, M. (Trans, eds.), Cambridge University Press, Cambridge, 2003;
- Keeley, M. (1981), "Organisations as non-persons" in *Journal of Value Enquiry*, 15, 1981, pp. 149-155;
- Lakoff, G. and Johnston, M. (1980), *Metaphors We Live By*, The University of Chicago Press, Chicago and London, 1980;
- Lewis, D. (1969), *Convention: A Philosophical Study*, Blackwell, Oxford, 2002;
- Mallon, R. Machery, E. Nichols, S. and Stich, S. (2009), "Against Arguments from Reference" in *Philosophy and Phenomenological Research*, 79(2), 2009, pp. 332-356;
- Nisbett, R. and Miyamoto, Y. (2005), "The Influence of Culture: Holistic versus Analytic Perception" in *TRENDS in Cognitive Science*, 9(10), 2005, pp. 467-473;
- Pettit, P. (1993), *The Common Mind: An Essay on Psychology, Society, and Politics*, Oxford University Press, New York and Oxford, 1993;
- Pettit, P. (1998), "Defining and Defending Social Holism", *Philosophical Explorations*, 3, 1998, pp. 169-184;
- Pinker, S. (1994), *The Language Instinct: The New Science of Language and Mind*, Penguin, New York and London, 1994;
- Quinton, A. (1975), "Social Objects", *Proceedings of the Aristotelian Society, New Series*, 76, 1975, pp. 1-27;
- Searle, J. (1995), *The Construction of Social Reality*, Penguin Books, London and New York, 1995;
- Sharrock, W. and Anderson, R. (1986), "Margaret Gilbert on the Social Nature of Language" in *Synthese*, 68, 1986, pp. 553-558;
- Sperber, D. and Wilson, D. (1995/1986), *Relevance: Communication and Cognition*, Second ed., Blackwell Publishing, Oxford UK and Cambridge MA;
- Weber, M. (1922), *Economy and Society*. Roth, G. and Wittich, C. (Trans, eds.), University of California Press, Berkeley, Los Angeles, California and London, UK, 1978.

ALESSANDRO SALICE

Universität Wien

alessandro.salice@univie.ac.at

SOCIAL ONTOLOGY AND IMMANENT REALISM

abstract

What are the intentional objects of groups' beliefs? This paper claims that they are immanent facts, i.e., facts which exist only within groups' minds. Since in relevant literature the notion of immanent object and the related theory of "immanent realism" arise in connection with the work of Franz Brentano, the paper begins by briefly sharing historical information on Brentano, making clear why – contrary to common belief – Brentano did not argue for immanent realism in his work. In a second part, I then look more closely at groups' beliefs and illustrate why the insight of immanent realism – despite its historically inadequate reconstruction – can bear on my initial question. In doing so, I pay particular attention to John Searle's theory of institutional facts, using it as a conceptual basis to develop my own pseudo-Brentanian approach. This approach allows me to introduce a further class of social entities in the last part of the paper: contrary to institutional facts the immanent entities of collective beliefs presuppose neither the assignment of functions nor the generation of deontologies, but they do presuppose groups' beliefs for their existence. Being the precipitates of collective experiences, such entities are intrinsically related with the first plural person perspective and hence play an important role in what we may call the "cultural layer" of social reality.

keywords

Social facts, immanent facts, cultural reality, social ontology, immanent realism

“[...] Dornröschen und Rotkäppchen sind Gestalten der deutschen Märchenwelt, die zur gemeinsamen Umwelt unseres Volkes gehört. Keinem von uns wird es einfallen zu behaupten, daß jeder sein eigenes Dornröschen habe. Es hat seine fest umrissenen Züge, seine persönliche Eigenart und Schicksale, an denen so wenig zu rütteln ist wie an denen einer historischen Persönlichkeit. Wir würden sehr energisch Verwahrung einlegen, wenn ihm jemand Züge andichten wollte, die ihm nicht zugehören. Unsere Märchenwelt nun hat eine gewisse Topyk, eine Eigenart, die sie von der anderer Völker, z.B. der chinesischen, charakteristisch unterscheidet.”
(Stein 1922, 134-135)

- 1. Introduction** In this paper I am mainly interested in the ontological status that the intentional objects of groups' beliefs bear. Suppose that I am a member of, e.g., a religious community and that I share with the members of this community the belief that, e.g., only someone with divine powers can be the leader of the community. In this scenario, what does it mean for a collective to share this belief? What is the object of the belief? On the one hand, the belief is false simply because nothing in the world makes it true. On the other hand, the non-existing fact that the leader has divine powers – if believed – produces some effects, e.g., the community respects him and/or only he has the right to perform given actions within the community.

One of the assumptions that drives my argumentation is that within social ontology there is a distinction to be drawn between sociality and culturality. Sociality comprehends all those entities, practices and dynamics which are intrinsically social and which are not culturally modifiable. In this sense, social notions like sincerity, victory, obligations, etc., seem to fall within the domain of sociality: they are not context-relative (see Smith 2003) or do not bear “essential properties” (as phenomenologists would prefer to say, see Salice 2012). On the contrary, culturality comprehends all those social entities which are intrinsic to a given culture and hence cannot exist outside or apart from this culture. In accordance with this idea, here I will claim that the objects of social and cultural beliefs belong to an ontological category of its own kind, i.e., they are the “immanent objects” of social beliefs, and that these objects partly constitute the domain of culturality.

The present article is organized as follows: since the notion of immanent object and the related theory of “immanent realism” arise in relevant literature in connection with the work of Franz Brentano, I begin (§1) by briefly sharing relevant historical information on Brentano, making clear why – contrary to common belief – Brentano did *not* argue for immanent realism in his work. In a second part (§3), I then look more closely at cultural beliefs and illustrate why the insight of immanent realism – despite its historically inadequate reconstruction – can bear on issues related to the main question raised above: what are the objects of collective beliefs? In doing so, I pay particular attention to John Searle's theory of social construction, using it as a conceptual basis to develop my own *pseudo*-Brentanian approach to cultural reality (§4).

The literature has employed the phrase “immanent realism” to characterize a phase in Brentano’s thought which more or less corresponds to the period in which Brentano published his groundbreaking monograph *Psychology from Empirical Standpoint* (1864). Among other merits, in this book Brentano reactivates the Aristotelian-medieval notion of intentionality and puts it at the center of philosophical debates. According to Roderick Chisholm and a number of other commentators, the intentionality thesis – as Brentano introduces it in that book – strictly taken entails two different theses. The first is a psychological thesis which states that reference to an object is the distinguishing element between psychic and physical phenomena. All psychic phenomena, and psychic phenomena alone, are *directed* towards something. The second is an *ontological* thesis which concerns “the nature of certain objects of thought and of other psychological attitudes”, i.e., it concerns the ontological status of that “something” to which the mind is related (see Chisholm 1967, 201).

2. Franz Brentano and Immanent Realism

According to this second thesis, for Brentano the object of an intentional act is an ontologically subjective entity which exists *only in the mind*. Brentano allegedly reached this insight when attempting to offer a solution to the problem raised by the so-called “objectless presentations”. Indeed, immanent realism can easily account for all those cases in which an act is directed towards a non-existing entity: the act of thinking of Pegasus shows no structural difference from the act of thinking of Barack Obama – in both cases the act of thinking is directed towards an object existing merely *in the mind*. With regard to the latter, however, it happens to be the case that this inner object corresponds to an actually existing object.

Recently, convincing evidence has been delivered that Brentano did *not* claim such a position. Scholars have presented several arguments to show that Brentano never accepted immanent realism (see Antonelli 2012, Sauer 2006). Among the reasons adduced by these authors, probably the most important is that Brentano’s position did not share the modern view of relations as *n*-place predicates or properties. *If* the intentional relation were a dyadic property, *then* the inference from the existence of the relation to the existence of its relata was indeed justified. However, Brentano denies that the intentional relation has to be analyzed in terms of a dyadic relation into which the subject *and* the object enter, and so he was not committed to the idea that all intentional acts have an object of a *sui generis* ontological status. In other words, for Brentano, it does not follow that if *a* thinks of *b*, then *both* relata, *a* and *b*, exist in whatever sense of the term “existence”.

Rather, Brentano understood relations in an Aristotelian sense, namely as monadic predicates. Accordingly, the fact that I think of Obama, is not constituted by my mind, the relation of thinking and Obama as object existing in the mind. This fact only includes my mind *and* the monadic property thinking-of-Obama and it obtains regardless of whether Obama exists or not (i.e., it does not entail the *existence* of an object in *my* mind). To put this point in Brentanian terms, I think of Obama and not of the thought-Obama, I hear the sound and not the heard-sound, I see the object and not the seen-object, etc. Obama, the sound, the object, etc., can be called – and are actually called by Brentano – the “immanent” objects of the act, but it should now be evident that this expression does *not* imply that these objects are *in the consciousness*.

To be sure, Brentano reserves an ontological *niche* to what he calls the act’s “correlates” (German: *Korrelate*) – that is, the aforementioned thought-Obama, heard-sound, seen-object etc. Such correlates are indeed parts of my mind, but they are *not* the objects of the act. But if so, then in what sense are these correlates *in my mind*? A clear answer to this question can be found in Brentano’s lectures on *Descriptive Psychology* given in Vienna during the 1880s and 1890s (see Brentano 1982). In these lectures Brentano reinforces a statement he already made in his *Psychologie* that all acts are conscious acts.

Not only are all acts directed towards an object, they are also directed towards themselves. Insofar as they occur, they simultaneously grasp themselves and hence are conscious. Suppose I am thinking of Obama, then at the same time I know that I am thinking of Obama. And the same is valid for all intentional acts: if I perceive or imagine Obama, I know that I perceive or imagine Obama. Looking now at this *second* intentional relation in which the act grasps itself, the act encounters, on the one hand, the act's component or mode (I know whether I am merely *imagining* Obama or if I *perceive* him, etc.). But, on the other, it also encounters Obama-*qua*-object of this act, that is, it also grasps the thought-Obama or the perceived-Obama. So the correlate does indeed play a role in Brentano's theory of intentionality, but only a secondary one, as it were.

But if Brentano did not advocate immanent realism, then questions arise as to why it is important to deal systematically with this theory and why it is relevant for social ontology. What follows should hopefully clarify the answers to both questions. In short, the idea is that the beliefs of individuals cannot create reality. It does not follow from the fact that individuals take something to be true, that this something has to exist in any sense. The same does not hold, however, for group beliefs. If a group believes something, then what the group believes does exist in a perspicuous sense. Of course, this reality is not that of brute facts; rather, what the group creates is a social or – probably more precisely – a cultural reality, and this is a reality which exists only *in* the group's mind.

3. Social Ontology and Immanent Realism

To illustrate this intuition, consider the following declarative sentence and assume that it expresses a belief:

(1) Akihito is the emperor of Japan

The sentence refers to a fact (an existing state of affairs) and, hence, it is true. But what is its truth-maker? To answer this question, one might do well to distinguish between external and internal points of view, depending on who holds the belief expressed in the sentence.

Following Searle's account (see Searle 1995, 2010), the sentence is true because it refers to an *institutional* fact. This is not an ontologically objective fact (such as the one that Mont Blanc is 4810m high), but an ontologically subjective one. If mankind were to disappear, then nations, emperors and, more generally, all institutional facts would disappear too. On the contrary, Mont Blanc would still remain.

What is then the structure of this institutional fact? What, in other words, are its constituents? First of all, sentence (1) suggests that the property being-emperor-of-Japan is exemplified by an individual, i.e., by Akihito. But this suggestion is false, for exemplification is an objective relation which characterizes objective facts, whereas the fact at issue is ontologically subjective. According to this analysis, (1) does not linguistically depict the state of affairs at issue adequately and hence has to be reformulated in the following way:

(1*) Akihito counts as the emperor of Japan

The "counting as" locution should convey the idea that Akihito is merely *held* as the emperor of Japan by someone. Still, the question now arises as to what kind of predicate the "counting as" locution expresses. In particular, is this a dyadic predicate (see Varzi 2007)? (Interestingly enough, here we come across a problem similar to that we discussed regarding Brentano's theory of intentionality.) If the "counting as" is a dyadic predicate and if sentence (1*) is true, then it would follow that two objects exist: on the one hand, Akihito and, on the other, the emperor of Japan.

But this view is not compatible with Searle's monist ontology. As he puts it regarding his preferred example of a 10 dollar bill: it is not the case that when I hold a 10 dollar bill in my hand, I hold two objects at the same time – the piece of paper and the dollar bill. One possibility to deal with this difficulty is to go for the idea that the “counting as” predicate is not primitive and has to be analyzed further. If x counts as y , then a given *status* – that is, a given *label* – is assigned by collective intentionality to x by means of an expressed or unexpressed speech act of declaration. The status or label is the y -term, and this status always goes with a certain function¹. Returning to the setting of our initial example, in Japan, if you are the the *Tennō*, you can fulfill certain functions that go with this status; for example, you are in charge for the convocation of the Diet. Since all status goes hand in hand with functions, Searle introduces the concept of “status function” to refer to all those functions which can be fulfilled only in relation to a given status. As a consequence, the fact at hand consists of the status function “emperor” imposed on an individual (i.e., Akihito) via collective intentionality.

There are two further necessary conditions that need to be fulfilled in order for this fact to exist. The first are the deontologies generated by the status function. As Searle puts it: “status functions = intentional facts \rightarrow deontologies” (Searle 2010, 23). The emperor has the right to adjourn the convocation of the Diet, but he also has the obligation to perform ceremonial functions. Without such deontologies, i.e., without the rights and obligations connected with the status, there are no institutional facts.

The second condition has to do with the beliefs that the group entertains. If the *creation* of this fact goes back to the imposition of a status to the x -term by means of a declaration, its “continued existence” is secured only if a group G (for example, for argument's sake, the Shinto community in Japan²) believes or accepts that Akihito is the emperor of Japan (Searle 1995, 117f). Here, again, if G does not believe/accept that Akihito is the emperor of Japan, the institutional fact at issue does not exist.

We have now found an answer to our initial question: to what state of affairs does the belief expressed by sentence (1) refer? The answer is: the belief is directed towards a complex entity (a status function assigned to an individual via collective intentionality) which exists only as long as the two aforementioned conditions are fulfilled. Since this state of affairs subsists, sentence (1) is true. Note, however, that the sentence analyzed so far expresses the belief of an *external* observer and that the state of affairs at hand is the intentional object of this particular belief. To make this clear, let's have a look at the second necessary condition, namely, at the group's beliefs.

What about the belief of group G ? To what fact does sentence (1) refer, if this sentence is taken to express the belief of G ? When G believes that Akihito is the emperor, G believes that Akihito has a divine nature. Since for G being-the-*Tennō* is co-referential with the property being an individual with divine nature (the senses of these predicates differ, but their objectual domain coincide), for G sentence (1) refers to the very same state of affairs as sentence (2):

(2) Akihito has divine nature

1 But in some cases you only have the label without the corresponding x : in 2010 Searle extends the domain of the notion of institutional fact by including the so-called “free-standing y -terms” (i.e., facts constituted only by a status function; e.g., corporations, electronic money, blindfold chess, etc.). Accordingly, the validity of the “ x counts as y ”-formula is now restricted to a limited number of institutional facts.

2 Only sociological research can ascertain whether or not there is a Shinto community in Japan which bears such belief, so this example may very well be fictional. Still, this does not seem to be relevant for the conceptual consequences to be drawn from it.

In other words, for G, “emperor” does not point primarily to a status function, but to a non-institutional concept. Hence, the belief of G and that of the external observer are *not* directed towards the same state of affairs. For the external observer defines the concept “emperor” only in terms of deontic powers and deontologies (see “[...] the [status] function is defined in terms of [deontic] power [...]”, Searle 1997, 451), whereas G defines this concept by virtue of properties which are taken to be objective. For G, to be the descendant of the Gods is not a socially construed fact: rather, the emperor is held to be the lineal descendent of the Gods by means of natural kinship (see Maraini 2003). Thus, the object of G’s belief is *not* an institutional, i.e., a socially construed, fact: the group does not believe that (1) is true by means of social agreement. Furthermore – under the plausible assumption that no man has divine nature – there is also no objective state of affairs to which G’s belief refers.

According to this last point, if we take (1) to express G’s belief, (1) is *false*. Still, this belief contributes to the constitution of an institutional fact. Let me stress at this point that it is not the belief as such that brings about the institutional fact, but this very *particular* belief, namely the belief *in* the fact that Akihito has divine nature. (The group may have thousands of other beliefs about Akihito, but only if it believes that Akihito has divine nature, is Akihito the emperor of Japan.) Accordingly, the sentence

(3) G believes that Akihito has divine nature

has to be in any case *true* for the corresponding institutional fact (expressed by (1*)) to exist. In other words, if sentence (3) were false (if G does not believe that Akihito has divine nature), then sentence (1*) would be false, too. To be sure, one could try to get rid of the linguistic reference to the non-existing fact enclosed in (3) by applying some sort of paraphrastic strategy. Following Russell’s suggestion, (3) could be paraphrased into:

(3*) There exists an *x* such that (i) G has a belief about *x*, (ii) *x* has divine nature and (iii) there is only one such *x*

But since *x* has no values, sentence (3*) is obviously false. Here I am not concerned with those cases in which *individuals* instead of collectives are mentioned in clause (i): it may very well be that the paraphrase works for *this* kind of case. Still, in the scenario at issue (3*) cannot be an adequate paraphrase of (3) since, as Searle points out, (3) captures a necessary condition for the existence of the corresponding institutional fact. Indeed institutional facts exist only in virtue of the beliefs of a community. But then, if (3) has to be true, what does G believe, when it believes that Akihito is the emperor?

4. Immanent Entities and the Cultural Layer of Social Reality

As suggested earlier in the paper, I contend that the object of G’s belief is a fact which (subjectively) exists only as long as the corresponding experience exists. This entity exists in the group’s mind and nowhere else and can be characterized as “immanent” in a *pseudo*-Brentanian terminology. Just as *Dornröschen* and *Rotkäppchen* have their own reality (their “definitely contoured traits [*seine fest umrissenen Züge*]”), but these figures existentially depend on the experiences of a group and *not* of that of an individual (in this case: on the fantasies of the German *Volk*), so the divine emperor of Japan (or more precisely: the fact that someone *is* the divine emperor of Japan) exists only in relation to the beliefs of a group.

This idea leads to an important consequence. Remember that immanent realism holds that immanent entities may or may not correspond to transcendent entities and that this does not make any structural difference for the belief of the individual. This idea can now be exploited in the case of

group beliefs: immanent entities can correspond to transcendent (institutional or natural) entities, but this does not make any difference for the group's experience. Nothing in the experience varies if its object does or does not exist. Let's discuss this point a bit further.

We have already dealt with the scenario in which the group believes that being-the-emperor and having-divine-nature are natural properties of Akihito. Here the immanent fact does not correspond to an objective or natural fact – for no such facts exist at all. But it does not *correspond* to an institutional fact either: although the institutional fact (subjectively) exists and although the immanent fact (together with the concomitant belief) is a necessary condition for its existence, G's belief is *not* concerned with this institutional fact. For what G intends, when it intends that Akihito is the emperor, is not what an external observer intends, when s/he intends that Akihito counts as the emperor. As we saw, sentence (1) is true if it expresses the belief of an external observer, but it is *false* if it expresses the belief of G.

So now one can modify this scenario by assuming that there is a group of social scientists (G') and that G' agrees with Searle's theory. Assume furthermore that G' knows that "emperor" is a status function and that it assigns this status function to Searle himself. In this example it happens to be the case that the immanent fact believed by G' *corresponds* to an institutional fact. In other words, the fact believed by G' corresponds to a given segment of institutional reality. As a consequence, the sentence

(4) Searle is the emperor

is true if we take it to express either the belief of an external observer or the belief of G'. Still, the immanent and the institutional facts at issue are and remain ontologically different. This last point also tells us that the existence of a transcendent fact (be it institutional or objective/natural³) does not make any structural difference for the belief of the group and that the group can be acquainted only with immanent entities. At this juncture, the central question raises as to how the term "belief" is employed here and whether the beliefs of individuals have the same features of group beliefs. I would answer this second question in the negative, i.e., beliefs of individuals *can* be acquainted with transcendent entities, but a discussion of this aspect would exceed the scope and purpose of this paper.

To conclude: in addition to institutional facts, in this paper I claimed that there is a further class of social entities which presuppose neither the assignment of functions nor the generation of deontologies, but which do presuppose collective beliefs (or group beliefs) for their existence. As these are merely the ontological precipitates of intentional experiences, as it were, such entities taken *per se* are not causally active. Only together with their corresponding beliefs do these facts have social effects. Being so intrinsically related with the first person plural perspective, they play an important role in what we may call the "cultural layer" of social reality. And indeed they constitute – at least an important domain of – this layer. To repeat the intuition that lead this analysis, we can say that the beliefs of individuals cannot create reality, but that of groups can. However, rigid ontological limits have to be set here as the only entities that these beliefs may create are "immanent" entities, i.e., entities which exist only in the group mind.

³ Obviously enough, the scenario at hand can be again modified by taking natural facts into account. If G believes that there is snow on the Mont Blanc, the immanent object of this belief corresponds to a natural fact if there actually is snow on the Mont Blanc.

REFERENCES

- Antonelli, M. (2012), “Franz Brentano’s Intentionality Thesis”, in A. Salice (ed.): *Intentionality*, Philosophia (Basic Philosophical Concepts), Munich, 2012;
- Salice, A. (2012), “Social Ontology as Embedded in the Tradition of Realist Phenomenology”, in M. Schmitz, H.B. Schmid et al. (eds): *The Background of Social Reality*, Springer, Dordrecht, 2012;
- Brentano, F. (¹1874, ²1924), *Psychologie vom empirischen Standpunkt*. With introduction and notes by O. Kraus, 2 vols., Meiner, Leipzig;
- Brentano, F. (1982), *Deskriptive Psychologie*, eds. by R. M. Chisholm and W. Baumgartner, Meiner, Hamburg, 1982;
- Chisholm, R. (1967), “Intentionality”, in P. Edwards (ed.): *Encyclopedia of Philosophy*, Vol. 4., Mcmillan, New York, 1967, 201-204.
- Maraini, F. (2003), *L’Agape Celeste. I riti di consacrazione del sovrano giapponese*, Luni Editrice, Milano, 2003.
- Sauer, W. (2006), “Die Einheit der Intentionalitätskonzeption bei Franz Brentano”, *Grazer Philosophische Studien* 73, 2006, 1-26;
- Searle, J. (1995), *The Construction of Social Reality*, The Free Press, New York, 1995;
- Searle, J. (1997), “Responses to Critics of *The Construction of Social Reality*”, *Philosophy and Phenomenological Research* 57-2, 1997, 449-458;
- Searle, J. (2010), *Making the Social World. The Structure of Human Civilization*, Oxford University Press, Oxford, 2010;
- Smith, B. (2003), “The Ontology of Social Reality”, *American Journal of Economics and Sociology*, Vol. 62-1, 2003, 285-299;
- Stein, E. (1922), “Beiträge zur philosophischen Begründung der Psychologie und der Geisteswissenschaften”, in E. Husserl (ed.), *Jahrbuch für Philosophie und phänomenologische Forschung*, Jg. 5, Niemeyer Verlag, Halle, pp. 1-284. New Edition: Stein, E., *Beiträge zur philosophischen Begründung der Psychologie und der Geisteswissenschaften* (edited by B. Beckmann-Zöller), in E. Stein, *Edith Stein Gesamtausgabe*, Vol. 6, Herder, Freiburg 2010;
- Varzi, A. (2007), *Il denaro è un’opera d’arte (o quasi)*. Ciclo di Conferenze e Seminari “L’uomo e il Denaro”. Quaderno 24. Università Cattolica del Sacro Cuore, Milano, 2007.

SESSION

2

SESSION 2

INSTITUTIONAL LIFE

Lorenzo Passerini Glazel (Università degli Studi di Milano-Bicocca)
Institutional Ontology as an Ontology of Types

Wojciech Żelaniec (University of Gdańsk)
On the Constitutive Force of Regulative Rules

*Guglielmo Feis (Università degli Studi di Milano),
Umberto M. Sconfienza (Tilburg University)*
Challenging the Constitutive Rules Inviolability Dogma

*Emanuele Bottazzi, Roberta Ferrario
(Istituto di Scienze e Tecnologie della Cognizione di Trento)*
Appearance Counting as Reality? Some Considerations on Stability
and Unpredictability in Social Institutions

Gaetano Albergo (Università degli Studi di Catania)
Does Ontogenesis of Social Ontology Start with Pretence?

LORENZO PASSERINI GLAZEL

Università degli Studi di Milano-Bicocca

lorenzo.passerini@unimib.it

INSTITUTIONAL ONTOLOGY AS AN ONTOLOGY OF TYPES

abstract

In present paper I investigate the relation between types and tokens, and its peculiar relevance for the ontology of institutional phenomena. I distinguish cognitive (or analogical) types and normative (or katalogical) types: while cognitive types are constructed a posteriori from analogies among pre-existing tokens, normative types are the prius of their tokens, tokens whose identity and effects are determined in the type.

Direction of fit, essential effects, and atypicalness in relation to normative types are investigated, and the hypothesis is ventured that the ontology of institutional phenomena (contrary to the ontology of natural phenomena) is primarily an ontology of types.

keywords

Social ontology, type/token relationship, direction of fit, atypicalness, institutional phenomena

The multiple impressions [ἐκτυπώματα εκτυπώματα] of a seal [σφραγίς sphragis] participate of their archetype seal [ἀρχέτυπος σφραγίς archetypos sphragis], and in every single impression the seal is present in its entirety and identity, without being part of any of its impressions.

Dionysius the Areopagite, De Divinis N ominibus, 644a, 7-9.

- 0. What are the Characteristic Circumstances in which We Postulate a Type?**
- In *Art and Its Object*, 1968, the British philosopher Richard Wollheim [London, 1923-2003] poses the following question: “What are the characteristic circumstances in which we postulate a type?” and he gives the following answer:
- A very important set of circumstances in which we postulate types [...] is where we can correlate a class of particulars with a piece of human invention: these particulars may then be regarded as tokens of a certain type. (Wollheim 1968, 94)¹*

Wollheim employs the “type vs. token” paradigm in his investigations on the ontology of works of art, so he thinks notably of some kinds of works of art, such as operas, ballets, poems, etchings, *etc.* Another important context in which we postulate types surely is the context of institutional phenomena, one of the main fields of investigation in social ontology. In present paper, I shall investigate some important aspects of the role of the “type vs. token” paradigm in social ontology and in the philosophy of institutional phenomena.

- 1. Types and Tokens**
- Investigations in social ontology and on the nature of institutional phenomena often resort to the “type vs. token” paradigm. It is, indeed, a fundamental cognitive process of the intellect to sort phenomena into different types; a process which is common to every field of knowledge. In social ontology, though, and particularly in the ontology of institutional phenomena, the type-token relation may play a different role from the mere *cognitive* one: it can play a specific *normative* role, which has a fundamental import not only at the *epistemological* level, but also (and maybe primarily) at the *ontological* level.

- 1.1. The “type vs. token” paradigm in Charles Sanders Peirce**
- In *Prolegomena to an Apology for Pragmaticism*, 1906, the American philosopher and semiotician Charles Sanders Peirce [Cambridge (Massachusetts), 1839-Milford (Pennsylvania), 1914] introduced a conceptual paradigm doomed to have a great relevance in semiotics and in investigations on social

¹ Wollheim sets here types against other generic entities, like classes and universals. For an analysis of the peculiarities of the concept of “type” as opposed to other generic concepts in categorization, see (Passerini Glazel 2005).

and institutional phenomena: the dyadic “type vs. token” paradigm.

Peirce elucidates the paradigm “type vs. token” through the example of word count:

A common mode of estimating the amount of matter in a manuscript or printed book is to count the number of words. There will ordinarily be about twenty the’s on a page, and of course they count as twenty words. In another sense of the word “word,” however, there is but one word “the” in the English language; and it is impossible that this word should lie visibly on a page or be heard in any voice, for the reason that it is not a Single thing or Single event. It does not exist; it only determines things that do exist. Such a definitely significant Form, I propose to term a Type. (Peirce 1960, 4.537, vol. IV, 423)

A “Type” thus, in Peirce’s definition, is a significant “Form” that “determines things that do exist”, though it doesn’t itself exist as a single thing, as an *ídion*².

The single actual concrete occurrences of a Type are named “Tokens” by Peirce:

A single event which happens once and whose identity is limited to that one happening or a Single object or thing which is in some single place at any one instant of time, such event or thing being significant only as occurring just when and where it does, such as this or that word on a single line of a single page of a single copy of a book, I will venture to call a Token. (Peirce 1960, 4.537, vol. IV, 423)

The relation between a type and its tokens is a relation of *instantiation*:

In order that a Type may be used, it has to be embodied in a Token which shall be a sign of the Type, and thereby of the object the Type signifies. I propose to call such a Token of a Type an Instance of the Type. Thus, there may be twenty Instances of the Type “the” on a page. (Peirce 1960, 4.537, vol. IV, 424)

It is quite obvious that, in attributing a predicate to an object, to an act, or to an event, we may refer *either* to the *type* of object (of act, of event), or to a single, concrete occurrence of that type of object (of act, of event), to a single *token* in its individuality.

To distinguish the attribution of a predicate to a *type* from the attribution of a predicate to a *token*, I propose to adopt Amedeo Giovanni Conte’s [*Pavia, 1934] “eidographic vs. idiographic” paradigm.

A predicate which is attributed to a *type qua type* (to a *type* as an *eídos* in Conte’s terms) is an *eidographic* predicate³.

A predicate which is attributed to a *token qua token* (to a *token* as an *ídion*) is an *idiographic* predicate⁴.

When I say, for instance:

[1] Dog is man’s best friend,

I am making reference to the *type* of animal “dog” in general (in comparison to other types of animals): this is an *eidographic* predication.

But when I say:

[2] The dog Jack is John’s best friend,

I am making reference to a single dog, Jack (in comparison to other dogs, or to other John’s friends): this is an *idiographic* predication.

² The adjective ἴδιος *idios* in ancient Greek means “proper, peculiar, particular, singular”.

³ The Greek word εἶδος *eidos* means “idea, form, species, type”.

⁴ “Eidographic” is a neologism of Conte’s; “idiographic” is borrowed by Conte from the German philosopher Wilhelm Windelband (“idiographisch”).

Similarly, when I say:

[3] In the game of rugby, a try is worth five points,

I am making reference to the (institutional) *type* of event “try”: this is an *eidographic* predication.

But when I say:

[4] In last rugby match England v. Ireland, the try scored by Ireland at the beginning of second half turned the tide of the game,

I am making reference to a single event, to a single *token* of the type of event “try”: this is an *idiographic* predication.

Again, when I say:

[5] In Italian law, in the absence of contrary stipulation, from marriage derives the community of property between the spouses,

I am making reference to the (institutional) *type* of act “marriage”: this is an *eidographic* predication.

But when I say:

[6] Andrea and Francesca’s marriage took place on September 29th, 2012,

I am making reference to a single marriage, to a single *token* of the type “marriage”: this is an *idiographic* predication.

1.3.
An Asymmetry
Between
Eidographic
and Idiographic
Predicates

Between *eidographic* and *idiographic* predicates there is an asymmetry.

Let’s consider examples [3], [4], and [5], [6], which are examples of institutional phenomena.

What is (eidographically) predicated of the *type* “try” in example [3] (the fact that a try is worth five points) can be predicated *salva veritate* of the *token* “try scored by Ireland etc.” (Ireland’s try surely was worth five points).

But the contrary doesn’t hold: what is (idiographically) predicated of the *token* “try scored by Ireland etc.” in example [4] (the fact that *that* particular try turned the tide of the game) cannot be predicated *salva veritate* of the *type* “try”.

Similarly, what is (eidographically) predicated of the *type* “marriage” in example [5] (the fact that from marriage derives the community of property in the absence of a contrary stipulation) can be predicated *salva veritate* of the *token* “Andrea and Francesca’s marriage”.

Again, the contrary doesn’t hold: what is (idiographically) predicated of the *token* “Andrea and Francesca’s marriage” in example [6] (that it took place on September 29th, 2012) cannot be predicated *salva veritate* of the *type* “marriage”.

The foregoing examples show that, on the one side, some essential *eidographic* properties of institutional types are necessarily “transmitted” to its tokens: a token of a try is necessarily worth five points, because the type “try” is conventionally stipulated to be worth five points; a token of marriage necessarily constitutes the community of property in the absence of a contrary stipulation, because that is stipulated for the type “marriage” in Italian law.

On the other side, *idiographic* properties of the tokens (some properties which belong to the tokens as *idia*) cannot (by definition) be transmitted to the type⁵.

⁵ An investigation on what kind of properties are necessarily transmitted, what kind of properties can be transmitted, what kind of properties cannot be transmitted between types and tokens (and in which direction) is beyond the purposes of this paper. A first investigation on this subject has been carried out by British philosopher Richard Wollheim: see (Wollheim 1968).

The same doesn't necessarily hold for natural types: the fact that we say that "Dog is man's best friend" depends on the fact that most dogs (most tokens of the type "dog") are friendly and useful to humans, and not vice versa.

Sub 1.2. (*Eidographic vs. idiographic predicates*) I have drawn a distinction between predicates which are attributed to a *type qua type* (to a type as an *eidos*) and predicates which are attributed to a *token qua token* (to a token as an *idion*).

But what is the relation existing between a *type* and its *tokens*?

Generally speaking, the relation existing between a type and its tokens is a relation of *conformity*: a *type* and its *tokens* must correspond one to the others, they must *fit* each other, to be, respectively, the *first*, the type of those tokens, the *second*, the tokens of that type.

But this relation of conformity may take up one of two opposite directions.

In some cases, it is the *type* that has to fit the tokens; in other cases, on the contrary, it is the *tokens* that have to fit the type.

To the relation existing between types and tokens, thus, applies the concept of "direction of fit", introduced by the American philosopher John R. Searle [*Denver, 1932] (and inspired to him by British philosopher G.E.M. Anscombe [Limerick, 1919-Cambridge, 2001])⁶.

Searle in (Searle 1975) distinguishes two directions of fit:

- (i) a "word-to-world" direction of fit,
- (ii) the opposite "world-to-word" direction of fit⁷.

A prefiguration of the concept of "direction of fit" is in Thomas Aquinas' [Roccasecca, 1225-Fossanova, 1274] *Summa Theologiae*:

Veritas consistit in adaequatione intellectus et rei [...]. Intellectus autem qui est causa rei, comparatur ad ipsam sicut regula et mensura, e converso autem est de intellectu qui accipit scientiam a rebus.

Quando igitur res sunt mensura et regula intellectus, veritas consistit in hoc, quod intellectus adaequatur rei, ut in nobis accidit, ex eo enim quod res est vel non est, opinio nostra et oratio vera vel falsa est.

Sed quando intellectus est regula vel mensura rerum, veritas consistit in hoc, quod res adaequantur intellectui, sicut dicitur artifex facere verum opus, quando concordat arti. Sicut autem se habent artificata ad artem, ita se habent opera iusta ad legem cui concordant. Iustitia igitur Dei, quae constituit ordinem in rebus conformem rationi sapientiae suae, quae est lex eius, convenienter veritas nominatur. Et sic etiam dicitur in nobis veritas iustitiae. (Aquinas, Summa Theologiae, part I, Question 21, art. 2)⁸

This passage from *Summa Theologiae* suggests promising connections between the idea of two opposite directions of fit and investigations on the type-token relation, on *de dicto* and *de re* truth, on "normative truth" and "truth of norms".

6 See (Anscombe 1957), (Searle 1975), (Searle 1985).

7 More recently, (Searle and Vanderveken 1985) considered cases (such as the case of declarations) in which there is a double direction of fit ("world-to-word-to-world"), and cases in which there is a "null or empty" direction of fit.

8 "Truth consists in a correspondence between the intellect and reality. Now, an intellect that is a cause of the relevant real thing is related to it as a rule and measure, whereas the converse holds in the case of an intellect that takes its knowledge from the thing. Thus, when, as happens with us, the things are the measure and rule of the intellect, then truth consists in the intellect's correspondence to the thing. For, it is because reality is (or, respectively, is not) such-and-such that our beliefs and statements are true (or, respectively, false). By contrast, when the intellect is the rule or measure of the things, then truth consists in the thing's correspondence to the intellect. So, for instance, the craftsman is said to produce a true work when that work agrees with his craft. Now just actions are related to the law with which they accord as artifacts are related to their craft. Therefore, God's justice, which establishes among things the order that conforms to the measure of His wisdom, i.e., His law, is appropriately called "truth". And this is why, even among us, one speaks of the truth of justice".

I'm not going to pursue all of these inspiring connections in present paper. I shall restrict here to the distinction of two kinds of types (*cognitive types*, *normative types*) according to their respective direction of fit.

2.1.
The Direction of
Fit of Cognitive
Types (or
Analogical Types)

2.1.1. In some cases, I said, it is the *type* that has to fit the tokens.

This is the case, for instance, with zoological types (such as “dog”, or “*canis lupus familiaris*”) and botanical types (such as “apple tree”, or “*malus domestica*”): more generally, this is the case with what I propose to call “*cognitive types*”.

Cognitive types have the function of reproducing (in the sense of representing) and synthesizing what the (typical) properties of the tokens of that type *are*.

It is because most singular tokens of the type “dog” are friendly and useful to men, that we use to say that “Dog is man’s best friend” (not the opposite).

Cognitive types thus have a “*type-to-world*” direction of fit: the type has to reproduce and reflect the world as it is, and so it is the *type* that has to fit the tokens.

2.1.2. If there is no conformity between type and tokens, it is the *type* that is *inadequate*.

Here is an example (a well-known example in modern and contemporary epistemology): a predicate of the type “swan” used to be, up to the end of XVII century: “Swan has white plumage”⁹.

The discovery of black swans in Australia at the end of XVII century led to rectify the type: now we know that “Swan can have white or black plumage”. (I guess nobody tried to “rectify” the colour of Australian swans.)¹⁰

2.1.3. As the examples of black swans shows, with regard to *cognitive types* we usually have what the Norwegian sociologist Johan Galtung [*Oslo, 1930] called *cognitive expectations*: i.e. expectations that we are prepared to revise in case of dissonance between our expectations and reality¹¹.

2.1.4. Another term I propose for cognitive types is “*analogical types*”: the type, indeed, is here (a *posteriori*) constructed through the recognition of the analogies existing among the tokens; the tokens are the *prius*, the type is the *posterius*¹².

2.2.
The Direction of
Fit of Normative
Types

2.2.1. In other cases, it is the *tokens* that have to fit the type.

This is the case, for instance, with conventional and institutional types (such as “try” in rugby, “marriage” in law, “10 euros banknote”, the word “the” in English), and, more generally, with what I propose to call *normative types*.

Normative types have the function of determining what something in the world *shall* be like to have a certain conventional or institutional value.

Through a normative type we do not *reproduce* what the properties of the tokens of the type *are*; on the contrary, we *determine* what the properties of the tokens *shall be*.

It is because the type “try” in rugby is conventionally and normatively stipulated to be worth five points, that we can say that the particular try scored by Ireland, was worth five points (and not the opposite).

The direction of fit of *normative types*, thus, is a *world-to-type* direction of fit: it is the tokens that have to fit the type, it is the tokens that have to reproduce (in the sense of replicating), to reflect the type.

⁹ Cfr. Juvenal’s famous verse: “*Rara avis in terris, nigroque simillima cygno*” (Juvenalis, *Saturae*, VI, 165).

¹⁰ I remind that in the gender *Cygnus*, besides white and black species, there is also a black-necked species, native to South-America.

¹¹ See (Galtung 1959).

¹² I set *analogical types* against *katalogical types*: see infra § 2.2.5.; my paradigm “*analogical types* vs. *katalogical types*” is inspired by the distinction between *ana-logical* and *kata-logical* in the thology of Hans Urs von Balthasar [Luzern, 1905-Basel, 1988]: see (von Balthasar 1985).

2.2.2. If there is no conformity between type and tokens, it is the *tokens* that are *inadequate*.

A try scored during half-time in a rugby match, or a try scored by a player in his own in-goal area, are no try at all, and no point will be awarded to any team.

2.2.3. With regard to *normative types* we usually have what Galtung called *normative expectations*: i.e. expectations that we are *not* prepared to revise in case of dissonance between our expectations and reality¹³.

2.2.4. That types (may) have a *normative* function with regard to tokens is an idea which is not alien to Peirce's semiotics. In a former formulation of the "type vs. token" paradigm, Peirce used, for "type", the term "*Legisign*" (defined as "a law that is a sign"), and, for "token", the term "*Replica*".

Here is Peirce's passage from *A Syllabus of Certain Topics of Logic*, 1903:

A Legisign is a law that is a Sign. This law is usually established by men. Every conventional sign is a legisign. It is not a single object, but a general type which, it has been agreed, shall be significant. Every legisign signifies through an instance of its application, which may be termed a Replica of it. Thus, the word "the" will usually occur from fifteen to twenty-five times on a page. It is in all these occurrences one and the same word, the same legisign. Each single instance of it is a replica. The replica is a sinsign. Thus, every legisign requires sinsigns. But these are not ordinary sinsigns, such as are peculiar occurrences that are regarded as significant. Nor would the replica be significant if it were not for the law which renders it so. (Peirce 1998, 2, 291)

2.2.5. Another term I propose for normative types is "*katalogical types*": the tokens are derived (they descend) from the type; the type is the *prius*, the tokens are the *posterius*.

In § 2. (*Two opposite directions of fit between a type and its tokens: cognitive types vs. normative types*) I distinguished *normative types* (*katalogical types*) from *cognitive types* (*analogical types*), and I adduced as examples of normative types some *conventional* and *institutional* types.

Conventional and institutional types of objects, of acts, of events, are among the main subjects of investigation in social ontology.

In present § 3., I shall focus on institutional types of acts, and I will underline a peculiarity of institutional types of acts, as opposed to non-institutional types of acts¹⁴.

Baptizing, entering into a sale contract, marrying, scoring a try, are all examples of institutional acts: acts that can be performed only within, and in virtue of, a particular institutional frame.

Running, opening a window, felling a tree are non-institutional actions: actions that can be performed independently of any institutional frame.

In following § 3.1. (*Extrinsic consequences of non-institutional acts*) and § 3.2. (*Intrinsic effects of institutional acts*), I will show an important difference existing between

- (i) the relation between the tokens of a non-institutional action and the type they are tokens of
- and
- (ii) the relation between the tokens of an institutional act and the type they are tokens of.

¹³ Galtung's "*cognitive expectations vs. normative expectations*" paradigm is fruitfully applied by Luhmann in his view of norms in sociological perspective: see (Luhmann 1969).

¹⁴ In philosophy of institutional phenomena, *institutional facts* are usually opposed to *brute facts*. Brute facts are facts that, in Searle's words, "can exist without human institutions"; on the contrary, *institutional facts* are facts that "require human institutions for their very existence".

3. *Intrinsic Effects of Institutional Acts vs. Extrinsic Consequences of Non-Institutional Acts*

3.1.1. *Non-institutional actions* like running, opening a window, felling a tree, may produce some *consequences*: whenever one runs, a translation of his body in space takes place; whenever one opens a window, the window gets open, and fresh air gets in; whenever one fells a tree, the tree falls on the ground. All of these consequences are produced by every particular action in its individual singularity, by virtue of mechanical cause-effect relations: all of these consequences are produced by every action in its *ídon*, independently of the fact that anybody recognizes them as *tokens* of a particular *type* of action.

3.1. *Extrinsic Consequences of Non-Institutional Actions*

3.1.2. All of these cause-effect consequences of non-institutional types of actions are *extrinsic consequences*: they are not inscribed in the *essence* of a type, or at least they are not *determined by the type*: they are merely and immediately actual at a token-level.

Cause-effect relations exist between *tokens*, between *ídon*, not between types: despite their universalization by induction in natural sciences, causal laws of nature “operate” at *token-level*, not at *type-level*.

3.2. *Intrinsic Effects of Institutional Acts*

3.2.1. *Institutional acts* like baptizing, entering into a sale contract, marrying, scoring a try, produce some specific *institutional effects*.

Whenever one baptizes somebody, whenever one enters into a sale contract with somebody, whenever one marries, whenever one scores a try, some institutional effects, specific to every type of act, are produced in the world.

Whenever a player scores a try in a rugby match, for instance, his team is awarded five points. This effect isn’t produced by virtue of mechanical cause-effect relations; it is produced by virtue of the fact that it is ascribed to (and inscribed into) the (normatively constituted) type of event “try” that it is worth five points.

Every token-try in rugby produces the effect of earning five points to the scoring team only by virtue of it being a token of the type “try” in rugby¹⁵.

As Maria-Elisabeth Conte [Soest in Westfalen, 1935 -Pavia, 1998] suggests, a token of an institutional act merely *triggers* [*löst aus*] the effects of the type; but the triggered effects are effects of the type, they are effects which are *intrinsic* to the type.

In other words, the effects of an institutional act are produced *through the tokens*, but *in virtue of the type*.

3.2.2. Unlike the extrinsic consequences of non-institutional acts, thus, the effects of an institutional type of act are *intrinsic effects* specific to every type of institutional act: these effects are ascribed to (and inscribed into) the *type*, the *eídon*, of every act.

In the world of institutional phenomena, institutional norms generally operate at a *type-level*, not at a token-level.

3.2.3. Adopting the lexicon of Scholasticism, the type of an institutional act is the (necessary) *causa prima* of the effects of its tokens, while the tokens are but a (contingent) *causa secunda* of those effects. To the relation between type as *causa prima* and tokens as *causa secunda* of the effects of an institutional act suit the following words of Aquinas (I^a-II^{ae} q. 19 a. 4 co):

Effectus plus dependet a causa prima quam a causa secunda, quia causa secunda non agit nisi in virtute primae causae.

The effect depends more on the first cause than on a secondary cause, since a secondary cause acts only in the power of the first cause.

15 See (Peirce 1903, 291): “Nor would the replica be significant if it were not for the [Legisign] law which renders it so.”

In § 3. (Intrinsic effects of *institutional acts* vs. extrinsic consequences of *non-institutional acts*) I pointed out that (unlike non-institutional acts) institutional acts produce their specific effects by virtue of their correspondence to a type.

Let's now focus on cases of dissonance, of non-conformity between type and tokens.

I propose to distinguish two different kinds of "non-conformity to a type", modelled on Aristotle logical distinction between *negation* (ἀπόφασις *apophasis*) and *privation* (στέρησις *steresis*)¹⁶:

- (i) a relative, and privative, non-conformity to a type, which I call "privative (or steretic) atypicalness";
- (ii) an absolute, and negative, non-conformity to a type, which I call "negative (or apophatic) atypicalness"¹⁷.

I call "*privative* (or "*steretic*", from Greek στερητικός *steretikos*, "privative") *atypicalness*" the atypicalness which consists in a *partial non-conformity* of a token to a type.

Two examples:

[7] The atypicalness of a lease contract in which something else from money is the consideration for lease, is *privative atypicalness*.

[8] The atypicalness of a counterfeit ten euros banknote, is *privative atypicalness*.

Both the lease contract and the counterfeit banknote are atypical (not because they are completely alien to a type, not because they don't have any type, but) because they partially deviate from their type, because they don't fully fit the type of which they are (atypical) tokens.

I call "*privative*" this kind of atypicalness because it consists in a (scalar, progressive) deficiency of typicalness (of conformity) to a type¹⁸.

I call "*negative* (or "*apophatic*", from Greek ἀποφατικός *apophatikos* "negative") *atypicalness*" the atypicalness which consists (not in a partial non-conformity to a type, but) in *irreducibility to any type*, in the absence of a corresponding type.

Three examples:

[9] The atypicalness of a 27 euros banknote is negative atypicalness.

[10] The atypicalness of castling in draughts is negative atypicalness.

[11] The atypicalness of scoring a try in football (soccer) is negative atypicalness¹⁹.

A 27 euros banknote, castling in draughts, scoring a try in football are atypical not because they deviate from a type, nor because they don't fully fit a type: they are atypical because there is no

¹⁶ See Aristotle, *Metaphysica*, 1056a 29.

¹⁷ See also (Passerini Glazel 2005, 223-228) and (Passerini Glazel 2012).

¹⁸ Privative atypicalness is privative as well as privative is the gnostic and neoplatonic conception of evil: evil is an imperfection, a deficiency of being, a lack of good.

¹⁹ The game of rugby is quite provocative for philosophers of institutional phenomena: according to a tradition, indeed, the game of rugby originated on November 1st, 1823, on the grass of Rugby Public School, when, according to a memorial tablet that lays there, William Webb Ellis "with a fine disregard for the rules of football as played in his time, took the ball in his arms and ran with it, thus originating the distinctive feature of the rugby game". It was properly by not fitting the rules and types of the game he was playing, that Webb Ellis originated a new game, with its new rules and types. It was the (negative) atypicalness of Webb Ellis act that inspired the types of a new game: the game of rugby.

4.
Two Kinds of
Atypicalness:
Privative
Atypicalness
vs. Negative
Atypicalness

4.1.
Privative
Atypicalness

4.2.
Negative
Atypicalness

corresponding type at all: there is no type “27 euros banknote” in European monetary system, there is no type “castling” in the game of draughts, there is no type “try” in the game of football. They are not tokens of a type, because there is no type, because a corresponding type doesn’t exist. Example [10] (castling in draughts) is inspired by the famous Ludwig Wittgenstein’s passage:

Statt “Man kann nicht im Damenspiel nicht rochieren”, sage: “Es gibt im Damenspiel kein Rochieren”.
(Wittgenstein, 1967, § 134, 315)²⁰

I call “*negative*” this kind of atypicalness because it consists in a (non-scalar, binary) *absence* of typicalness: there is no typicalness because there is no type²¹.

4.3.
The Opposite
Presuppositions of
Privative Atypicalness
and Negative
Atypicalness

Privative atypicalness and *negative* atypicalness have opposite presuppositions. *Privative* atypicalness, consisting in a *partial deviation* of a token from a type, presupposes the *presence* of a reference type (with regard to which privative atypicalness of a token is predicated). The type “ten euros banknote”, for instance, is presupposed by a counterfeit ten euros banknote. On the contrary, *negative* atypicalness, consisting in *irreducibility* to any type, presupposes the *absence* of a reference type.

4.4.
Three Kinds
of Normative
Impossibility

The paradigm “*privative atypicalness vs. negative atypicalness*” illuminates (and is itself illuminated by) the phenomenon of *normative impossibility*. Normative impossibility is any form of rule-related impossibility (of *nomophoric* impossibility, in Amedeo Giovanni Conte’s and Paolo Di Lucia’s [*Milano, 1966] lexicon)²². Normative impossibility (rule-related impossibility) may be the result of three different normative situations.

- (iii) A first kind of normative impossibility derives from the presence of deontic rule of prohibition: it is normatively impossible, in Italy, to smoke in public premises because there is a legal norm that prohibits to smoke in public premises.
- (iv) A second kind of normative impossibility derives from the presence of an anankastic-constitutive rule that poses a necessary condition that cannot be fulfilled: it is normatively impossible for a 40-year old person to participate to a competitive state exam if an anankastic-constitutive rule poses as a necessary condition to enter the state exam being younger than 35.
- (v) A third kind of normative impossibility derives from the absence of an eidetic-constitutive rule that constitutes a particular type of act or event: it is normatively impossible to castle in draughts because there is no constitutive rule of draughts game that constitutes the type of move “castling”.

In the game of chess, it is *possible* to castle: it is possible in virtue of the rules of the game; the possibility of castling is thus a *normative possibility*.

On the contrary, in the game of draughts, it is *impossible* to castle: it is impossible

- (i) not because of the presence of a deontic rule that forbids it,
- (i) not because of the presence of an anankastic-constitutive rule that poses a necessary condition that cannot be fulfilled;

²⁰ Not: “One can’t castle in draughts” but - “There is no castling in draughts”. Wittgenstein’s example is quoted as an example of “absence-impossibility” (impossibility due to the absence of certain rules) by Amedeo Giovanni Conte and Paolo Di Lucia in (Conte and Di Lucia 2012, 138). I think that my paradigm “privative atypicalness vs. negative atypicalness” may be fruitfully confronted with Conte and Di Lucia’s “presence-impossibility vs. absence-impossibility” paradigm, as well as with Conte and Di Lucia’s “nomophoric impossibility vs. non-nomophoric impossibility” paradigm.

²¹ Negative atypicalness is negative as well as negative is the Manichean conception of evil: evil is other than good, it is absence of good.

²² See (Conte and Di Lucia 2012, 136-137).

(i) but because of the absence of a rule: the absence of an eidetic-constitutive rule that constitutes the type of move “castling”, and so makes it possible to instantiate an infinite number of tokens of “castling” in draughts. It is impossible, here, to produce a token of a type that doesn’t exist.

The *third* kind of normative impossibility (the normative impossibility to castle in draughts), which is seemingly the most radical kind of normative impossibility, is a normative impossibility due to the *absence* of a type, due to *negative atypicalness*²³.

The impossibility due to the absence of a type recalls Stanisław Jerzy Lec’s words:

Many things didn’t come to existence due to the impossibility to give them a name.

5.1. In the foregoing §§ 1., 2., 3., 4. of this paper I showed:

- (ii) that some essential *eidographic properties* of institutional types are necessarily transmitted from a type to its tokens;
- (iii) that conventional, institutional, and, in general, normative types have a *world-to-type direction of fit*;
- (iv) that the effects of an institutional act are produced *through the tokens*, but *in virtue of the type*;
- (v) that there are cases where some institutional phenomenon is *impossible due to the absence of a corresponding type*.

5.2. These considerations suggest the idea that *types* in social ontology, and notably with regard to institutional phenomena, do not play a merely *cognitive* role, but a *normative* one, and they arise a question: Can there exist institutional phenomena without types? In other terms: Can there exist an institutional reality at all without types?

5.3. This *normative* role of institutional types has an import not merely at the *epistemological* level, but chiefly at the *ontological* one.

While in natural phenomena a token is what it is independently of the type it is ascribed to, in social ontology a token of an institutional phenomenon is what it is only in virtue of the type which it is an instance of.

Types, in social ontology and notably with regard to institutional phenomena, seem to be not only condition of *conceivability*, but also condition of *possibility* of their tokens.

Reformulating an idea formulated by the Italian legal philosopher Amedeo Giovanni Conte about norms, types are not merely *epistemologically* transcendental: they are also *ontologically* transcendental with regard to their tokens²⁴.

The hypothesis I formulate here is that the ontology of institutional phenomena is primarily an *ontology of types*, while the ontology of natural phenomena is primarily an *ontology of tokens*²⁵.

5.
Ontologies
of Tokens vs.
Ontologies of
Types

²³ Conte and Di Lucia call this impossibility “*absence-impossibility*” (impossibility due to the *absence* of a certain rule), as opposed to “*presence-impossibility*” (which is due to the *presence* of a certain rule): see (Conte and Di Lucia 2012, 138).

²⁴ See (Conte 1962) and (Conte and Di Lucia 2012).

²⁵ A similar suggestion has been (provocatively) envisaged by the Scottish legal philosopher sir Neil MacCormick: MacCormick states that “at least the world of legal institutions is a world safe for Platonists” (MacCormick 1986, 55).

REFERENCES

- Anscombe, G.E.M. (1957, ²1963), *Intention*, Cornell University Press, Ithaca (New York), 1957, ²1963;
- Anscombe, G.E.M. (1958), "On Brute Facts", *Analysis*, 18, 1958, pp. 69-72;
- Azzoni, G. M. (1988), *Il concetto di condizione nella tipologia delle regole*, CEDAM, Padova, 1988;
- Balthasar, H. U. von (1985), *Thologik. Zweiter Band. Wahrheit Gottes*, Johannes Verlag, Einsiedeln, 1985;
- English translation: *Theo-logic. II: Truth of God*, San Francisco, Ignatius, 2004; Italian translation by Guido Sommovilla: *TeoLOGICA. II. Verità di Dio*, Jaca Book, Milano, 1990;
- Benoist, J. and Kervégan, J. F. (eds.) (2008), *Adolf Reinach. Entre droit et phénoménologie*, CNRS Éditions, Paris, 2008;
- Betti, E. (1971), *Interpretazione della legge e degli atti giuridici*, Giuffrè, Milano, 1971;
- Conte, A. G. (1962), *Saggio sulla completezza degli ordinamenti giuridici*, Giappichelli, Torino, 1962. Also in Amedeo Giovanni Conte, *Filosofia dell'ordinamento normativo*, Giappichelli, Torino, 1997, pp. 71-302;
- Conte, A. G. (1986), "Fenomeni di fenomeni", in Giuseppe Galli, ed., *Interpretazione ed epistemologia. Atti del VII Colloquio sulla interpretazione (Macerata 1985)*, Marietti, Torino, 1986, pp. 167-198. Second edition (with variants): *Rivista internazionale di Filosofia del diritto*, 63 (1986), pp. 29-57. Third edition in Amedeo Giovanni Conte (1995), *Filosofia del linguaggio normativo. II. Studi 1982-1994*, Giappichelli, Torino, pp. 313-346;
- Conte, A. G. (1988), *Eidos. "An Essay on Constitutive Rules"*, in Giuliano di Bernardo (ed.), *Normative Structures of the Social World*, Rodopi, Amsterdam, 1988, pp. 251-257;
- Conte, A. G. (2001), "Dimensions of Nomic Freedom", in Ian Carter/Mario Ricciardi (eds.), *Freedom, Power, and Political Morality. Essays for Felix Oppenheim*, Palgrave, New York, 2001, pp. 69-78;
- Conte, A. G., Di Lucia, P. (2012), *Adýnaton. Four Dychotomies for a Philosophy of Impossibility, Phenomenology and Mind*, 2, 2012 pp. 134-143;
- De Monticelli, R.(2003), *L'ordine del cuore*, Garzanti, Milano, 2003;
- De Nova, G. (1974), *Il tipo contrattuale*, CEDAM, Padova, 1974;
- De Vecchi, F. (ed.) (2012), *Eidetica del diritto e ontologia sociale. Il realismo di Adolf Reinach*, Mimesis, Milano, 2012;
- Di Lucia, P. (1997) *L'universale della promessa*, Giuffrè, Milano, 1997;
- Di Lucia, P. (2003a), *Normatività. Diritto linguaggio azione*, Giappichelli, Torino, 2003;
- Di Lucia, P. (2003b), "Ontologia del dovere eidetico", in Paolo Di Lucia, *Normatività. Diritto linguaggio azione*, Giappichelli, Torino, 2003, pp. 141-162;
- Di Lucia, P. (2003c), "Tre modelli dell'ontologia sociale", in Paolo Di Lucia (ed.), *Ontologia sociale. Potere deontico e regole costitutive*, Quodlibet, Macerata, 2003, pp. 9-24;
- Di Lucia, P. (2003d), "Tre specie di dovere eidetico. Un'analisi ontologica", in Paolo Di Lucia (ed.), *Ontologia sociale. Potere deontico e regole costitutive*, Quodlibet, Macerata, 2003, pp. 217-235;
- Di Lucia, P. (ed.) (2003e), *Ontologia sociale. Potere deontico e regole costitutive*, Quodlibet, Macerata, 2003;
- Galtung, J. (1959), "Expectations and Interaction Processes", *Inquiry*, 2, 1959, pp. 213-234;
- Gomperz, H. (1939), *Interpretation. Logical Analysis of a Method of Historical Research*, van Stockum & Zoon, Den Haag/The University of Chicago Press, Chicago, 1939;
- Husserl, E. (1913), *Ideen zu einer reinen Phänomenologie und phänomenologische Forschung. Erstes Buch: Allgemeine Einführung in die reine Phänomenologie*, in "Jahrbuch für Philosophie und phänomenologische Forschung", 1 (1913), pp. 1-323 (Max Niemeyer, Halle).
- Husserl, E. (1938), *Erfahrung und Urteil. Untersuchungen zur Genealogie der Logik*, Academia, Prag, 1938; new edition: *Klassen*, Hamburg, 1948;
- Husserl, E. (1950), *Cartesianische Meditationen und Pariser Vorträge*, edited by S. Strasser, in *Husserliana*, I, Nijhoff, Den Haag, 1950;
- Lorini, G. (2000), *Dimensioni giuridiche dell'istituzionale*, CEDAM, Padova, 2000;
- Lorini, G. (2002), "L'atto giuridico nella filosofia dell'atto", in Giuseppe Lorini (ed.), *Atto giuridico*, Adriatica, Bari, 2002, pp. I-XVII;

- Lorini, G. (2007a), “Atto thetico in Czesław Znamierowski”, in Amedeo G. Conte, Paolo Di Lucia, Antonio Incampo, Giuseppe Lorini and Wojciech Żelaniec, *Ricerche di Filosofia del diritto*. A cura di Lorenzo Passerini Glazel. Giappichelli, Torino, 2007, pp. 224-233;
- Lorini, G. (2007b), “Norma costruttiva in Czesław Znamierowski”, in Amedeo G. Conte, Paolo Di Lucia, Antonio Incampo, Giuseppe Lorini and Wojciech Żelaniec, *Ricerche di Filosofia del diritto*. A cura di Lorenzo Passerini Glazel. Giappichelli, Torino, 2007, pp. 79-88;
- Luhmann, N. (1969), “Normen in soziologischer Perspektive”, in *Soziale Welt*, 20 (1969), pp. 28-48. Partial Italian translation by Edoardo Fittipaldi and Lorenzo Passerini Glazel: *Aspettative e norme*, in Giuseppe Lorini and Lorenzo Passerini Glazel (eds.), *Filosofie della norma*, Giappichelli, Torino, 2012, pp. 283-298;
- MacCormick, N. (1986), “Law as Institutional Fact”, in Neil MacCormick and Ota Weinberger (eds.), *An Institutional Theory of Law*, Reidel, Dordrecht, 1986, pp. 49-75;
- Passerini Glazel, L. (2003), “Fitting Types. Tipi di atti e atti quali tipi”, in Paolo Di Lucia (ed.), *Ontologia sociale. Potere deontico e regole costitutive*, Quodlibet, Macerata, 2003, pp. 351-371;
- Passerini Glazel, L. (2005), *La forza normativa del tipo. Pragmatica dell'atto giuridico e teoria della categorizzazione*. Quodlibet, Macerata, 2005;
- Passerini Glazel, L. (2008), “Eidologique et eidonomique dans la philosophie de l'acte juridique: Adolf Reinach et Czesław Znamierowski”, in Jocelyn Benoist and Jean-François Kervégan (eds.), *Adolf Reinach. Entre droit et phénoménologie*, CNRS Éditions, Paris, 2008, pp. 59-69;
- Passerini Glazel, L. (2012), *Atto norma tipo. Tra pragmatica e ontologia del diritto*. Roma, Aracne, 2012;
- Peirce, C. S. (1931-1958), *Collected Papers of Charles Sanders Peirce*, 8 volumes, vol. 1-6, edited by Charles Hartshorne and Paul Weiss, volumes 7-8, edited by Arthur W. Burks, Harvard University Press, Cambridge (Massachusetts), voll. III e IV, 1960;
- Peirce, C. S. (1980), *Semiotica*. Testi scelti e introdotti da Massimo Achille Bonfantini, Letizia Grassi, Roberto Grazia. Einaudi, Torino, 1980;
- Peirce, C. S. (1989), *La logica degli eventi*. Edizione italiana a cura di Rossella Fabbrichesi Leo. Spirali, Milano, 1989;
- Reinach, A. (1913), *Die apriorischen Grundlagen des bürgerlichen Rechtes*, in “Jahrbuch für Philosophie und phänomenologische Forschung”, 1 (1913), pp. 685-847. Second edition, under the title: *Zur Phänomenologie des Rechts. Die apriorischen Grundlagen des bürgerlichen Rechts*, Kösel, München, 1953. Third edition, under the title: *Die apriorischen Grundlagen des bürgerlichen Rechtes*, in Adolf Reinach, *Sämtliche Werke. Kritische Ausgabe und Kommentar*, Philosophia, München, 1989, pp. 141-278. Italian translation by Daniela Falcioni: *I fondamenti a priori del diritto civile*, Giuffrè, Milano, 1990;
- Reinach, A. (1989), *Sämtliche Werke. Kritische Ausgabe und Kommentar*, Philosophia, München, 1989;
- Sacco, R. (1995), *La nozione di contratto*, in *Trattato di diritto privato* diretto da Pietro Rescigno. Vol. X. *Obbligazioni e contratti. Tomo secondo*, UTET, Torino, 1995, pp. 5-20;
- Searle, J. R. (1969), *Speech Acts. An Essay in the Philosophy of Language*, Cambridge University Press, London, 1969. Italian translation by Giorgio Raimondo Cardona: *Atti linguistici. Un saggio di filosofia del linguaggio*, Boringhieri, Torino, 1976;
- Searle, J. R. (1975), “A Taxonomy of Illocutionary Acts”, in K. Gunderson (ed.), *Language, Mind, and Knowledge*, University of Minnesota Press, (Minneapolis), 1975, pp. 344-369;
- Searle, J. R. (1983), *Intentionality: An Essay in the Philosophy of Mind*, Cambridge University Press, Cambridge, 1983. Italian translation by Daniele Barbieri: *Della intenzionalità. Un saggio di filosofia della conoscenza*. Bompiani, Milano, 1985;
- Searle, J. R. (1985), *Expression and Meaning: Studies in the Theory of Speech Acts*, Cambridge University Press, Cambridge, 1985;
- Searle, J. R. (1995), *The Construction of Social Reality*, The Free Press, New York, 1995. Italian translation by Andrea Bosco: *La costruzione della realtà sociale*, Comunità, Milano. New Italian edition: Torino, Einaudi, 2006;

- Searle, J. R., Vanderveken, D. (1985), *Foundations of Illocutionary Logic*, Cambridge University Press, Cambridge, 1985;
- Wittgenstein, L. (1981), *Zettel*, Blackwell, Oxford, 1981;
- Wollheim, R. A. (1968), *Art and its Objects. An Introduction to Aesthetics*. Harper & Row, New York. New edition: *Art and its Objects*, Pelican, 1968. Italian translation by Enzo De Lellis: *Introduzione all'estetica*, Isedi, Milano, 1974. New English edition with six supplemental essays: *Art and its Objects.v* Cambridge, 1980;
- Żełaniec, W. (2003), "Sull'idea stessa di regola costitutiva", in Paolo Di Lucia (ed.), *Ontologia sociale. Potere deontico e regole costitutive*, Quodlibet, Macerata, 2003, pp. 155-179;
- Żełaniec, W. (2007), "Regola costitutiva", in Amedeo G. Conte, Paolo Di Lucia, Antonio Incampo, Giuseppe Lorini and Wojciech Żełaniec, *Ricerche di Filosofia del diritto*. A cura di Lorenzo Passerini Glazel. Giappichelli, Torino, 2007, pp. 36-47.
- Znamierowski, C. (1921), "O przedmiocie i fakcie społecznym [Social objects and social facts]", in *Przegląd Filozoficzny*, 24 (1921), pp. 1-33;
- Znamierowski, C. (1924), *Podstawowe pojęcia teorii prawa. I. Układ prawny i norma prawna [Concepts fundamentals de la théorie du droit. Structure juridique et norme juridique]*, Fiszer i Majewski, Poznań, 1924;
- Znamierowski, C. (2006), "Norma costruttiva ed atto thetico", in *Rivista internazionale di filosofia del diritto*, 83 (2006), a cura di Giuseppe Lorini, pp. 283-285.

Wojciech Żelaniec
University of Gdańsk
zelaniec@aol.com

ON THE CONSTITUTIVE FORCE OF REGULATIVE RULES

abstract

The paper deals with the problem of delimitation of regulative and constitutive rules. I argue that while regulative rules are and remain a genus of their own, they do have a constitutive import, because they (i) define new forms of behaviour (behaviour compliant with them), (ii) take constituted entities as “input”, (iii) redefine old forms of behaviour as such to which a deontic modality applies and (iv) help to carve out new possibilities of behaviour from a continuum of only potentially distinguishable possibilities.

keywords

Norm, rule, constitutive, regulative, deontics

In a not yet so remote past, it was an unwritten rule of decorum on Polish public transport means to offer one's seat to an elderly or handicapped person standing in the aisle. Now no longer: you occasionally see whole rows of seats occupied by teenagers or persons in their twenties with sexagenarians or septuagenarians humbly standing in the aisle. Bend over to one of the young sitters and suggest to her the idea of giving up her seat to this lady or that gentleman and you will see eyes resplendent with innocent perplexity and a facial expression, not of a defiant "who are you to tell me that?" or "why should I?" but of sheer wonder at the very idea. Apparently, the idea has never crossed the young person's mind¹, strikes her as utterly eccentric, and does not exist to her as a human possibility, a *Lebensform*, in the Wittgensteinian² nor Thomas-Mannian³ sense.

But let us begin from the beginning.

Constitutive rules are often defined by oppositions to other kinds of rules. Amongst these oppositions that to "regulative rules" (that is, rules that say that someone must, must not, need not, may or may not, do or abstain from this or that) is probably the best-known. Searle has, in part, defined constitutive rules as in so far distinct from the regulative ones as they, as he put it, do not just regulate but also define new forms of behaviour⁴. On the face of it, the distinction seems clear-cut. (Except that it is not clear whether the distinction be ex- or, much rather, merely intensional⁵.)

Yet, on reflection one easily realises that regulative rules do, them too, define new forms of behaviour—namely, behaviour *compliant with them*. For instance: "Smoking aboard of this aircraft is prohibited": clearly, while not smoking as a form of outer behaviour is just not smoking, whatever its motive, taken jointly with its true motive—if, that is, the desire to comply with the norm in question is its true motive—it is different from, for instance, not smoking for lack of desire to smoke.

1 To be fair to Poland's younger generations, different behaviour is still observable, every now and again... Yet, it is typically persons over forty who vacate seats for those over sixty or the handicapped.

2 On "form of life" in Wittgenstein and in general (including Thomas Mann and the concept's philological history), see (Conte 1995a, 317f. footnotes 4-6).

3 (Mann 1926).

4 (Searle 1970, 33). Actually, Searle says "create or define", as if this were the same. This sloppiness does not remain without consequences, as I shall show later on.

5 Searle himself suggested (Searle 1970, 186f.) that in a sense even the Decalogue is a set of constitutive rules. I shall return to this issue later.

The Italian philosopher Gaetano Carcaterra has suggested to speak of a “constitutive force” of a norm (rule)⁶, rather than of “constitutive norms” (rules) *per se*. This is close to the idea that all norms partake of constitutivity to a degree. The border-line between the two kinds of norms may then come out blurred, or in any event not so clear-cut. Or maybe it is clear-cut, but only conceptually (intensionally), not extensionally? Or perhaps the matter is even more complex?

This essay is an investigation into this issue.

I presuppose this premise (which I consider incontrovertible): motives (true, as distinct from pretended ones, I stress) of a behaviour are part of the behaviour’s identity, that is, if they change, a different behaviour results. A person abstaining from smoking for a while is not “doing the same thing” whether she acts out of respect for the law⁷ (in this case, a norm which prohibits smoking at the person’s spatio-temporal location) or whether she simply has no wish to smoke (e.g. as a habitual non-smoker).

As long as only the outward aspect of the behaviour is considered, both actions (actually, omissions) look identical, but this is an illusion. Depending on what the motives are, the behaviour can take different courses and often does. For example, once the “no smoking!” rule has been abrogated, the temporarily non-smoking habitual smoker may well light a cigarette, while the one who has no desire to smoke will not in the least be affected. Also, one who observes rules out of a desire to observe them may have a further “pharisaic” desire, namely, to be seen as one who observes rules, and consequently his behaviour may be ostentatious or pronouncedly conspicuous, whereas a person who has other motives to behave as the rules say he or she should cannot have a desire like that. This is perhaps not frequent with omissions but think of the dignified immobility of one who refuses to cross a street on red at a crossing where most everybody is crossing the street regardless of the lights. Generally, depending of what our motives are, we shall react differently to new stimuli and modify our behaviour in a different way.

It will, perhaps, be responded that from the point of view of whoever lays down regulative rules it makes no difference what motives those who comply with them are guided by. Even if not smoking out of respect for the law is a new form of behaviour, it is, from the point of view of the legislator, equivalent to the old forms, even though not strictly speaking identical with them.

While this is, in a sense, true, it is also true that the legislator is not, and cannot be, quite indifferent to the motives of those whose behaviour she attempts to regulate. If she had been, she would not have cared to lay down any rules, hoping that people will behave the way she wants them to “for whatever motive”. Yet, she knows only too well that this won’t work: among “whatever motives” people happen to have there are, alas, great many which push them to behave in an undesirable way, for instance, to smoke on aeroplanes. The legislator knows this, so she provides a motive to behave the way she wishes, namely, the desire to comply with a rule, and this rule is precisely the regulative rule “Smoking is not permitted aboard of this aircraft”. Saying that the legislator does not aim at provoking a kind of behaviour “out of respect for the law”⁸ but a class of behaviours agreeing in the outward aspect only but done for “whatever motives” is nearly as adequate as saying that a buyer does not really will buying the commodity she is buying but only wills “acquiring” it (in “whatever way”)—as if the purchasing act were performed unbeknownst to her and without her consent. It is

6 (Carcaterra 1979). On p. 137 he speaks of a “*sostanza costitutiva*” of certain rules in a sense close to “being constitutive purely and simply”.

7 The Kantian “*Achtung fürs Gesetz*”, *Groundwork for the metaphysics of morals*, Akademie-Ausgabe IV, 400. Or out of fear of the punishment; in any case, a behaviour done “in function of” the law (Conte 2000).

8 See footnote 2.

true that a buyer is interested primarily in the object or service she is buying but it is also true that the immediate object of her willing is the act of purchase, *qua* the preferred way of acquiring the good or service.

Yet still, it is, in a sense, true that the new forms of behaviour that regulative rules give rise to are less interesting to those who draw up and enact these rules than are the new forms of behaviour constitutive rules give rise to in the eyes of their respective legislators. To the former, it is only essential that people behave as the rules say they ought to; they are ultimately interested in the “legality”, to speak with Kant⁹, of behaviour, just like a buyer is ultimately interested in acquiring the object and would welcome being given it for free. To the latter, and also to the “custodians” of constitutive rules (referees, etc.), it is, too, crucial that people should behave in conformity with constitutive rules for the sake of, among other things¹⁰, such conformity. A person who, due to her mimicking talents, or by sheer coincidence, behaves like everybody else during a religious ceremony or a session of a secret organisation, without knowing the rules, and without the intention to follow them, will not be considered to be “playing the game” and will be treated with distrust¹¹.

This is perhaps the most salient respect in which constitute and regulative rules differ. Nonetheless, because they give rise to a new form of behaviour in the sense here set out, regulative rules can be said to have a *constitutive force*.

However, even if in the final analysis it be conceded that regulative rules do, them too, define new forms of behaviour, it can be objected that these new forms of behaviour are not as radically new as the ones defined by constitutive rules. (Which, too, makes it unlikely, for most classical examples of constitutive rules, that anyone should behave in conformity with them while not knowing them.) How new is not smoking on board of an aeroplane for norm-related reasons? Clearly, it is not radically new: there is (and has been for long) the well-defined (outward) activity of not smoking, there is the deontic modality “you must not”, very well known too, the two combined form a rule, there is the human ability to do something because of a rule (either for the Kantian *Achtung fürs Gesetz*, or for fear of punishment) and that is it, it seems. The new form of behaviour is new in the sense in which a sentence which nobody has cared to formulate is new once it *has been* formulated, but no more: old words, old grammar, new sentence. Perhaps the sentence “the windows were open as the warmth had come”¹² has never before been formed in the English language, yet obviously it is just known words laid out in a known array, thus new in a rather weak sense.

Behaviour conforming to constitutive rules, by contrast, is not just so weakly new. In them, that which the deontic operator (if any) is attached to is not understood apart from the rules. A good example is “castling” in chess. Chess rules forbid castling if the king has already moved. But as distinct from forbidding smoking on board of a plane, the forbidden action is unintelligible (we do not know what it is) unless certain other rules specify its essence, and those rules are, again, constitutive. They *constitute* castling¹³.

9 *Metaphysics of morals*, Akademie-Ausgabe VI, 219.

10 Though, in most cases, not exclusively.

11 This presupposes, obviously, that *somebody* knows the rules. Which is why (as Guglielmo Feis M.A. of Milan University has brought to my awareness) in occasions where empty rituals are practiced and (next to) no-one remembers the underlying rules, there is a general uneasy feeling and an equally general tendency to treat one another with distrust. The Reader may pick his/her favourite example.

12 It is, actually, a translation of a Swedish sentence from a short story by Strindberg: “*Fönstren stodo öppna, ty värmen hade kommit*” (Strindberg 1913, 23).

13 It is remarkable that in this sense the syntax of the verb “to constitute” is different from what it is in non-constitutive-rules-related contexts, where it is basically an elaborate or “fancy” variant of “to be”. For instance in “This European project will constitute [sic] a significant breakthrough” (https://www.zsi.at/attach/PlakatA0_neu-1.pdf), where “will constitute a breakthrough” means just “will be a breakthrough”. Constitutive rules, in contrast, that constitute (in the sense relevant here) something-or-other (e.g. castling) are not that thing. (Rules of castling *are not* castling.) There is an “*obiectum constitutum*” or “*effectum*”, to use a word fittingly invented by Maria-Elisabeth Conte, see (Conte 1995, 282f., footnote 38).

It is crucial, here, that the definition of castling is given in the same set of rules as that which contains the prohibition just mentioned. Many regulative rules forbid, command or allow actions which are not immediately understood, but are defined and explained (constituted) somewhere else, *not* within the same body of rules.

For instance, the Ten Commandments would seem to be a classical example of a system of regulative rules. The commandment “Thou shalt not kill”¹⁴ is a clearly regulative rule since it forbids a type of action which we understand immediately, without need for any “constitution”¹⁵. But the commandment “Thou shalt not commit adultery”¹⁶ is not so immediately understood. What is “adultery” (the root “אָפֵּן” (nʔp))? Clearly, it is not simply every sexual activity whatever; it is a particularly “constituted” form of sexual activity, it can even be extended to inward activity of the mind¹⁷. Yet what matters is that it is not defined in the same body of rules, the Decalogue, as distinct from “castling”, which is explained in the same body of rules as that in which it is, in certain circumstances, prohibited, viz. the rules of chess¹⁸. This sets this commandment apart from typical examples of constitutive rules.

Arguing in this way, however, one runs the risk of making what must seem rather arbitrary decisions. Certainly, the concept of adultery is not explained in the text of the Decalogue itself¹⁹, but it is explained in the Bible (e.g., in the famous story of king David and Bath-Sheba, Uriah’s wife²⁰), of which (both formulations of) the Decalogue are a part. Why consider only a part of a text, rather than the whole, or rather this part than that part? It would be unwise, it seems, to expect that every system of rules should be as compact and surveyable as the rules of chess. But on the other hand it will be awkward to regard the whole of the Bible, and perhaps other Jewish and Christian texts, as one huge formulation of the Decalogue, and after all, some limits must be drawn... . So the Commandments may, with some plausibility, be regarded as a set of regulative rules, even though they presuppose a set of constitutive rules, which are given elsewhere (viz. in the Bible less Ex 20, 1-17 and Deut 20, 4-21).

To give another, perhaps more tractable, example: the German constitution (*Grundgesetz*, Basic Law) assigns various duties and rights to the “German” (*Deutscher*) in a series of what must seem (at least in its first part) undeniably regulative rules. For instance, it assigns to every German the right of freely choosing one’s profession²¹. But who or what is a “German”? There is, certainly, an intuitive concept of a German person, which implies such traits as being of German parents, having lived in Germany since one’s birth, holding German citizenship, speaking German as a mother-language, or a combination of such-like. Yet this intuitive concept is insufficient for the purpose of defining the subject of all the duties and rights laid down in a Constitution; therefore, the German Basic Law defines, in article 116 (1), the concept of a German as: ... *a person who possesses German citizenship or who has been admitted to the territory of the German Reich within the boundaries of 31 December 1937 as a refugee or expellee of German ethnic origin or as the spouse or descendant of such person*. Here, it is not possible to say that the regulative rules of which the German constitution in part consists presuppose constitutive rules specified *elsewhere*.

14 The fifth or sixth, depending on which division of the Commandments you employ.

15 Abstraction made from the question whether “to kill” is really the best translation of the Hebrew “חָצַר” (rʕh), rather than “to murder” or something else.

16 The sixth or seventh.

17 Mt 5, 28.

18 Art. 3, 8.b.(1).a of the FIDE Laws of Chess, <http://www.fide.com/fide/handbook.html?id=124&view=article>.

19 Ex 20, 1-17, Deut 20, 4-21.

20 2 Sam 11.

21 Art. 12 (1). Compare the view ascribed to a former prime minister of Italy, Mr. Silvio Berlusconi: “a working class kid should stay working class” (*il figlio dell’operaio deve fare l’operaio*), [http://www.flcgil.it/rassegna-stampa/nazionale/unita-
evviva-il-figlio-dell-operaio.flc](http://www.flcgil.it/rassegna-stampa/nazionale/unita-
evviva-il-figlio-dell-operaio.flc).

Thus, it must be said that regulative rules turn out to have, in this case not a constitutive force, but a *constitutive support*, in the sense that they take as “input” entities constituted by constitutive rules on an earlier stage or a deeper-lying level of the same system of rules.

At the latest here, though, the reader might start thinking that I have, from the outset, left out of account an important point, without which all attempt at delimiting regulative and constitutive rules is bound to remain fruitless. This point is the presence (in regulative rules) and the absence (in constitutive rules) of deontic operators, such as “may”, “must”, “ought to”, “has the right/duty to”, “is prohibited”, “may”, “may not” and the like. In fact, many prominent constitutive rules do not contain such operators. Also, because of this difference, the way of giving rise to behaviour, whether old or new, would for the two different kinds of rules be quite different. Regulative rules would give rise directly to—causally create—actual behaviour (if anyone cares to observe them), while constitutive rules would only define possibilities. This is because constitutive rules are grammatically descriptive sentences and while they need not describe anything that has actually occurred, they “describe” possible behaviour.

These are well-placed objections. A third one that could be added is that regulative rules with their deontic operators cannot be (easily) squeezed into the Searlean Procrustean bed of the canonical form of all constitutive rules, “X counts as Y in context Z”²².

Yet, as research has convincingly shown, it is not true that rules containing deontic operators cannot be constitutive²³. Amedeo Conte has distinguished a whole class of what he has called “deontic constitutive rules”²⁴. The mentioned rule restricting castling to situations in which the king has not yet moved is one of them²⁵. Also, all deontic sentences with operators like “may (not)” or “must (not)” can be understood as so-called “spurious deontic sentences”²⁶, i. e. descriptive propositions saying that a regulative rule is in vigour; then, regulative rules would be interpretable as constituting certain forms of behaviour as an *object* of duty, or other deontic modality, with respect to which they are not posterior. Searle is right²⁷ in claiming that (while stealing as such is old,) stealing as something you must not do is (comparatively)²⁸ new, and is created or defined by “Thou shalt not steal”²⁹. I think it is quite common human experience to be amazed on learning that what we thought to be a neutral activity we have innocently indulged in is a bearer (in the given normative context) of a deontic modality; most typically: is prohibited. For instance, this writer was once dumbfounded to find out that taking photographs of consenting adults was prohibited in the main building of the Faculty of Social Sciences of the University of Gdańsk. “Why should anyone issue a prohibition like that?” That is: why has this new form of behaviour: “photographing consenting adults in the said premises as prohibited” come into being? This is yet another way in which regulative rules may be said to have a *constitutive force*.

There seems then to be a perfect analogy between (constitutive) deontic rules governing the constituted entity called “castling”, and (regulative) deontic rules governing the equally constituted entity called “a German” (*Deutscher*), in the rules of chess and the German Constitution, respectively.

22 (Searle 1970, 35).

23 The view that they cannot is sometimes met with in popular discussions, the reason being the presumed Humean gap between “ought” and “is”. This is ironic, since the very distinction between the two kinds of rules was introduced by Searle for the purpose of bridging that gap (Searle 1964).

24 (Conte 1989, 243).

25 (Conte 1989, 244).

26 (Åqvist 2002, 154).

27 See footnote 4.

28 In the perspective of the entire history of humanity the proclamation of the Decalogue is a recent event.

29 The seventh or eighth Commandment.

Why not call them both “constitutive”? Why insist that the various rights and duties of a German are given in rules which, although they have a constitutive support, are themselves regulative?

Because, a reply would go, that kind of deontic modality as that governing castling is essential part of what castling is. If this should seem of little convincing force (as, after all, castling is what it is, i.e. the king and a rook switching fields, all restrictions come afterwards): the chess-bishop’s “duty” to move diagonally, the prohibition of touching the ball with one’s hands in association football (soccer) are quite distinctly essential to what the chess-bishop, resp. the game of football, is.

But is the prohibition of adultery not equally essential to a pious Jew or Christian? Is the freedom of choice of profession or occupation not essential to a (Federal Republic’s) German? The correct answer is, I propose, a “yes” when the pious Jew/Christian, or the German are considered *qua* such, and a “no”, when they are considered *qua* human beings. A pious Jew or Christian stops being such while committing adultery, yet obviously does not—human, all too human—stop being a human being. A modern German who believed to be legally bound to choose a profession or occupation and acted upon this belief would not immediately lose his citizenship but would count as seriously out of tune with his status as a German citizen³⁰. As a human being such a person, by contrast, would be intact.

Yet analogously, it will be objected, a chess-bishop moved horizontally or vertically would stop being a bishop, all the while remaining what it is outside of the chess context, viz. a piece of wood. A human chess-player who moved his bishop in this way would, for this moment, stop being a chess-player, all the while remaining a human being.

There is an important difference, however: The status of a chess-bishop lent to a piece of wood, that of a chess-player or a footballer lent to a human being, is not directly derivative of any essential trait of either. Pieces of wood have no natural propensity or inclination to become counters in a board game, and human beings have no such inclination to sports. There are human beings, to be sure, who do have such an inclination but there also are some who do not, and there is nothing pathological or otherwise abnormal in not liking sports. All (healthy and not handicapped) human beings have, by contrast, a propensity to, and a need of, living in a “polity”, an organised community of other human beings³¹. It is not, of course, essential to human beings as such to be citizens of the Federal Republic of Germany but it is essential to them to have a status (be it that of a citizen or subject, or the like) in a republic, or a kingdom, or a duchy or some other form of a state-like entity. As Kant once noted³², even a people of demons would need a state³³. Similarly for adultery and Jewish/Christian piety: it is not essential (at least from a non-believer’s or agnostic’s perspective) to have that religion or any religion at all, but it is essential to have an ethics (that is, to lead an organised, disciplined life) and again for an ethics it is essential to regulate the matters of sexual life in one way or another. What is condemnable “adultery” from the point of view of one ethical system may be permissible or even highly laudable “polyamory” from another’s, but some “line” on, and some structuring for, sexual love is necessary.

Thus, systems of regulative rules do contain constitutive elements, such as rules pertaining to “a German” or “adultery”, yet they are set apart from systems of “normal” constitutive rules because the constituted elements are articulations of basic traits and needs of human nature³⁴. Of course, such

³⁰ Candidates to German citizenship are obliged to pass a test of, among other things, their knowledge of the German Basic Law. A person like that would not pass this test (if carried out seriously).

³¹ Aristotle’s “πολιτικὸν ζῷον” (“*politikon zoon*”, “social animal”), *Politics* I, 1253a, and III, 1278b.

³² *Perpetual peace*, Akademie-Ausgabe VIII, 366.

³³ As reminds us (Cortina Orts 2003, 16).

³⁴ As set forth by Cotta in his (Cotta 1991). See my (Żełaniec 2011). Or sometimes of non-human nature, such as in the definition of a “built-up area” in the Highway Code.

articulations contain and cannot help containing, given the limitations of the human mind and will, a large amount of arbitrariness³⁵; hence different forms of polity and different ethical systems. Yet they all contrast with systems of constitutive rules (which, in their turn, have regulative elements), because these latter give form to accidental propensities of human beings, and define non-essential (though often interesting and useful) entities³⁶.

And finally, there is yet another way in which regulative rules can be weakly constitutive or have a constitutive force. Consider the young sitters on Polish trams and buses. That which more traditionally-minded passengers think is their obligation (viz. vacating their seats for the benefit of the elderly and handicapped) does not exist to them as a possibility to contemplate even though it is conceptually very simple and not at all “constituted”. Yet it is not obvious, in a way in which occupying a free seat during a ride is, it does not suggest itself to your tired body as “the most natural thing to do”—quite on the contrary. It has to be “carved out” from a continuum of most diverse possibilities (to vacate a seat just for a stop, for one minute, for some other time, getting someone else to vacate hers, waiting till the elderly person asks you to be allowed to rest on your seat³⁷, making that person to sit on your knees, and so on) none of which is obvious and none of which suggests itself to you on a merely psycho-physiological basis. And exactly that “carving out” (*constituting in a weak sense*) was done by the rule of decorum I mentioned at the beginning of this paper—when it was still in vigour.

35 Which is why my position is not simply a variant of *jusnaturalism*.

36 This does not apply to the Kelsenian Basic Rule, if it be a constitutive rule, that is. See (Azzoni 1988, 71-78).

37 Which, as Guglielmo Feis M.A. of Milan tells me, is the Italian method, at least amongst the young.

REFERENCES

- Åqvist, L. (2002), "Deontic logic", in Gabbay, Dov M., F. Guenther (eds.), (2002), *Handbook of philosophical logic*, Vol. 8, Kluwer, Dordrecht, 2002, pp. 147-264;
- Azzoni, G. M. (1988), *Il concetto di condizione nella tipologia delle regole*, CEDAM, Padova, 1988;
- Carcattera, G. (1979), *La forza costitutiva delle norme*, Bulzoni, Roma, 1979;
- Conte, A. G. (1989), "Regola costitutiva in Wittgenstein", in Conte, Amedeo G. (ed.), (1989), *Filosofia del linguaggio normativo. I. Studi 1965-1981*, Giappichelli, Torino, 1989, pp. 237-254;
- Conte, A. G. (1995), "Paradigmi d'analisi della regola in Wittgenstein", in Conte, Amedeo G. (ed.), (1995), *Filosofia del linguaggio normativo. II. Studi 1982-1994. Con una lettera di Norberto Bobbio*, 1995, pp. 265-312;
- Conte, A. G. (1995a), "Fenomeni di fenomeni", in Conte, Amedeo G. (ed.), (1995), *Filosofia del linguaggio normativo. II. Studi 1982-1994. Con una lettera di Norberto Bobbio*, 1995, pp. 313-346;
- Conte, A. G. (2000), "Nomotropismo: agire in funzione di regole", *Sociologia del diritto*. Vol. 27 n. 1, pp. 7-32;
- Cortina Orts, A. (2003), *Covenant and contract*, Peeters, Leuven, 2003;
- Cotta, S. (1991), *Il diritto nell'esistenza*, Giuffrè, Milano, 1991;
- Mann, T. (1926), *Lübeck als geistige Lebensform*, Quitzow, Lübeck, 1926;
- Searle, J. R. (1964), "How to Derive 'Ought' from 'Is'", *The Philosophical Review*. Vol. 73, pp. 43-58;
- Searle, J. R. (1970), *Speech acts*, Cambridge University Press, Cambridge, 1970;
- Strindberg, A. (1913), "Nybyggnad", in Strindberg, A., *Samlade skrifter. Femtonde delen. Utopier i verkligheten*, Albert Bonniers Förlag, Stockholm, 1913, pp. 11-92;
- Żelaniec, W. (2011), "Cotta on the ubiquity of the legal", in *Annali della Facoltà di Giurisprudenza di Taranto Anno IV*, Cacucci, Bari, 2011, pp. 765-777.

GUGLIELMO FEIS
Università degli Studi di Milano
guglielmo.feis@unimi.it

UMBERTO M. SCONFIENZA
Tilburg University
u.m.sconfienza@tilburguniversity.edu

CHALLENGING THE CONSTITUTIVE RULES INVIOIABILITY DOGMA

abstract

According to Searle (1969) a constitutive rule creates the very possibility of the entity it rules. Later on, other scholars - such as Benoist (2003), Conte (1983), Guastini (1983) - pointed out the inviolability of constitutive rules.

In this paper we want to challenge the inviolability dogma by asking the following question: is it possible to have violations of constitutive rules? We will argue that a parametrical approach to constitutive rules can do away with the inviolability dogma. As a result, our approach allows for violations and introduces two different ways of exiting a game (semantic exit vs. practical exit).

keywords

Constitutive rules, inviolability dogma, social practice rigidity

Introduction This paper concerns the problem of the inviolability of constitutive rules (CRs). The problem of constitutive rules inviolability is the unavoidable fallout of what we call the standard theory of constitutive rules¹.

Several rhetorical devices are usually employed in order to justify the standard theory and its fallout against the criticisms that is possible to move against them to the extent that is possible to deem “dogmatist” the position of those who employ those devices. This paper has two main aims: firstly, it tries to debunk the dogmatist position by showing the ontological problems that such a position entails; secondly, it tries to put forward an alternative theory of constitutive rules that does away with the inviolability fallout.

This paper has 4 sections: in section 1, we introduce the main concepts and theories that we are going to deal with: constitutive rules, the inviolability dogma and some criticism and rejoinders on both side of the debate (in support and against the dogmatist position); in section 2, we challenge the dogmatist position by sketching our parametrical theory and we try to understand whether or not our parametrical theory of constitutive rules can replace the standard theory of constitutive rules; in section 3, by comparing the two theories, we analyze some of the positive consequences of our proposal (i.e. when and how violations occur and a theoretically meaningful new division between different ways of exiting a game). A conclusion in which we acknowledge the possible criticisms that such a new approach may raise will follow in section 4.

- 1. Constitutive Rules and Inviolability: The Inviolability Dogma (ID)** By inviolability dogma (ID) we mean the thesis according to which changing something regarding a constitutive rule is not a violation of the rule but the creation of another rule-constituted type that has no relation with the first type.

¹ We will call standard theory the way of approaching the constitutivity issue along the line of Searle (1969) according to which constitutive rules create the very possibility of new behaviors - as in the rules of chess. The Italian legal philosopher Amedeo G. Conte exploited a full and complex taxonomy out of the Searlean and Rawlsian - see Rawls (1955) - intuition of constitutive rules. His main idea was to investigate CRs both as conditions of possibilities and as conditions of conceivability of the objects they rule. He further articulated his taxonomy working out different kind of conditions. See his volumes of collected papers Conte (1989, 1995, 2001).

We want to show that the issue of violation is twofold: on the one hand it is possible to violate the codified constitutive rule of a game by way of modification (think about playing soccer 12 vs. 12 instead of 11 vs. 11); on the other hand it is possible to go too far away and, instead of modifying a game, changes go too far and we exit the game (we will distinguish a semantic and a practical way of exiting a game).

We claim, against the ID, that violations as modifications occur and that, when a theory of CRs reconstructs two games as unrelated types - i.e. the classic move of the standard theory to deny that CRs' violations happen - there is a much more complicated story to dig into that concerns different ways exiting a game².

The debate whether constitutive rules are inviolable or not is a consequence of two different characteristics of constitutive rules. On the one side CRs are, as the name itself suggests, rules, i.e. they display a prescriptive element that tells how a certain entity, object or person, has to act or behave. On the other side, CRs are constitutive, i.e. the object ruled by the constitutive rule essentially depends on it and cannot be shaped otherwise.

As *rules*, CRs, prescribe a certain behavior as something that ought to be followed as opposed to something that ought to be avoided. This fact entails, at least *prima facie*, that there could have been a different stipulation according to which the constituted object X could have been characterized by property P1 instead of property P2. In this sense - i.e. as rules in general - CRs, are breakable because once a certain property is chosen, you prevent the ruled object from having some other properties that are not the ones you are ascribing to it through *that* CR.

As *constitutive*, CRs are not ordinary rules, like “do not smoke” or “drive on the left side of the road”, indeed, they are sort of stipulative definitions, they do not only rule over a certain entity but they create that very entity out of nothing. Take for example the CR “the bishop moves only diagonally”, this CR does not only tell us how the bishop should move, it tells us also how the bishop should move *in order to be a bishop*. In this sense, then, CRs cannot be broken because not following them means failing to create the very entity the rule is supposed to rule over.

The inviolability dogma poses major challenges; all the more if we use CRs in order to explain rule-based social practices or if we want to explain social practices by means of reference to the rules that constitute and regulate them. Constitutive rules are constitutive as a whole - i.e. it is the totality of the rules that is constitutive of a practice - so, different sets of rules end up constituting different practices. If a practice like “giving a lecture as plenary speaker” is determined by inviolable constitutive rules, then we end up with a practice that is too rigid, i.e. every minor departure from the constitutive rule results in a practice that is no longer “giving a lecture as plenary speaker”. Imagine that one of the CRs that creates the practice “giving a lecture as plenary speaker” is that the speaker has to lecture in front of her audience. Assume now that the speaker, for the whole duration of the talk, speaks from behind the audience: given this situation, would we say that she is still “giving a lecture as a plenary speaker”? Probably not. However, if she goes talking from behind the audience only from time to time (maybe because she is a nervous walking-speaker), then our perception would be different.

Closely related to the rigidity of a practice, comes the problem of *the multiplication of the practices that are too rigid*. Given the link between the set of CRs and the object constituted, we

² It is far beyond our scope to state that there was an *Ur*-game that has been modified several times in order to create all the actual games we may experience nowadays.

1.2 Two Facets of Constitutive Rules

1.3. Inviolability as a Source of Rigidity for Social Practices

are forced to claim that there are a lot of different unrelated types for all the spectrum of the very similar lecturing practices.

1.3.1. Single Rule Constitutivity vs. Constitutivity of the System as a Whole

Here we will refer to single CR as “constitutive”. Nonetheless, in the theory of CRs, this is somehow controversial because, properly speaking, CRs are constitutive only in a certain context of rules, i.e. it is the system of rules that is constitutive as a whole.

Taking this observation seriously will involve sorting out whether there are rules whose constituted object is more or less dependent on the rest of the system in which the rule is embedded. This research has never been carried out. Take the two following constitutive rules inside their system: “castling” in the game of chess and the article that constitute the Italian flag (art. 12 of the Italian Constitution) in the Italian constitution. We think that, if we analyze the dependence of those single rules within the rest of the system, castling appears more connected to the rest of the system than the Italian flag.

Note as well that this assumption of “constitutivity relativized to systems” dramatically diminishes any presupposed constitutive powers of CRs and decreases the possibility of the existence of anything such as eidetic-constitutive rules (i.e. rules that defines the essence of a rules-dependent object): essences are not supposed to be related to a system of rules!

Summing up: a dogmatic CRs theory cannot reject our proposal using a strict “constitutivity relativized to systems” approach because the same approach, literally taken, will demolish many of the achievements of the CRs theory.

1.4. Cheating: An Ad Hoc Counterexample to the ID?

The most intuitive reply to the inviolability dogma is the observation that you can actually cheat in CRs-based game. Take the abused example of chess: what if I move the bishop as a queen and you do not recognize my move? Can we say that there was a violation of a CR?

The standard reply to that from the dogmatist will run as follows: “*Strictly speaking*”, the inviolabilist says, “we are not playing chess when cheating occurs but something else” - for example (Ross 1968, 54)³ - without any detail on what this “something else” is. Further, as we have seen, this reply has compromising ontological consequences related to type multiplication.

Despite this, we think the following reply makes a good point against the *strictly speaking* mysticism: suppose we are playing chess, I try to cheat and you spot me. You will probably say that what I am doing is something illicit related to a CR and not that I am trying to move into a different game (without your consent)⁴.

Another way to back up the *strictly speaking* reply is to develop a theory of what the “something else” is. Amedeo G. Conte developed the concept of nomotrophism or acting-with-reference-to-rules⁵. The idea is that, even when we cheat, we are guided by some (others) CRs, especially the rules of the game constituting the victory in that game, i.e. the rules that makes it worth for us to try to cheat.

The issue of nomotrophism is really interesting but it fails to prove that a violation of a constitutive

³ This passage is quoted by (Roversi 2007, 101) as well. He then observes that (Ross 1968, 24) contradicts himself by claiming that “to cheat in any game is to make a ‘move’ which violates the constitutive rules of the game, with the intention that the other players, unaware of the violation, should take it as a regular move”.

⁴ This sort of reply is to be found in (Żelaniec 2003, 164-165).

⁵ Conte’s first work on the topic is Conte (2000).

rule has never occurred: even if we admit that I am cheating because of some other CRs, this fact is not enough to say that there was no violation of the CR I violate in order to have an act of cheating.

It seems, from the examples of the previous section, that the thesis according to which CRs are inviolable has some major flaws that not only jeopardize our ontology (type explosion) but also our grasping of what really goes on in the social practices (rigidity)⁶.

We need to take seriously into account the somehow paradoxical question: are there breakable constitutive rules?

Our proposal is then to do without the ID and elaborate a different theory of constitutive rules according to which it is possible to violate CRs. Our theory will revolve around the concept of parameter.

Considering that CRs are the products of human acts, and therefore, subjected to human creativity, discussion and agreement, we think it is likely that there will be more ways to implement CRs.

Intuitively speaking, for the time being, it is possible to play soccer even if we are not 11 vs. 11, the different ways can be sorted out through a variable in which different value can be entered. Some values (12 vs. 12) will comply with the CR performing its task, some will not (0 vs. 0)⁷.

Here we sketch how the parametrical constitutive rule theory works in order to deal with the problems of the standard theory outlined above.

The main argument runs as follows⁸:

1. In the standard theory, every value not complying with the range set by the parameter is “out of the game” (it defines another game) – and this is a problem, because in this way games and social practices are too rigid and too many. The standard theory sees an excess of conceptual economy using only types and tokens. We would like to add a third element to the theory: a structure derived from the recognition of the elements of the game disciplined through parameters.
2. We should rather conceive the possibility of violating the rules while standing “within the game”.
3. Hence, let us introduce the new concept of a parametrical constitutive rules that, through different values that can be assigned to a parameter, enables a lot of different practices which would seem deviant to be accounted as standing within the game because they are in accordance with the parameter and hence with the rule.
4. Violations here arise only when (i) we are not playing according to the codified CR and yet (ii) we do not fall outside the range of possible values for a certain parameter.
5. Outside the range of values the game changes. However, this is not a problem anymore for

⁶ In a forthcoming paper we tackle in more details all the problems of the CRs standard theory and give more details about our parametric approach. Here we want to go further showing how this approach can be used against the ID, enabling us to have breakable CRs.

⁷ We think that our proposal can be seen as the first step of a full development of Snyder’s intuition according to whom “The notion of somebody breaking the rules constantly but yet playing chess does not make sense. We can make no sense of the notion of a chess player who *always* makes illegal moves. (But, of course, *occasional failure* to observe the rules is not a ground for disqualification as a player)”. See (Snyder 1971, 171, emphasis added).

⁸ For a better development of our parametrical theory see our forthcoming paper *Structuring constitutive rules: a parametric shift. Creating a new game vs. Modifying a game*.

2. Do We Really Need the Constitutive Rules Inviolability Dogma?

2.1. A Sketch of a Parametrical Theory of Constitutive Rules

6. two reasons: (i) when we fall outside a game, according to the parametrical CRs theory, it will be a different game and not only a variation of a previous game starting from its codified CRs; (ii) given the empirical imprinting of our theory, we are nonetheless able to cope also with the historical evolution and the comparison of different games rather than considering them different unrelated tokens as it was in the standard theory.

Now consider the rule (officially codified by FIFA) related to the number of players (call it “N° CR”):

(N° CR) soccer is played 11 vs. 11.

Now, we try to isolate its parameter. From (N° CR) we obtain:

(*N° CR) soccer is played X vs. X.

In (*N° CR) the variable X stands for a range of possible values according to which we are still allowed to say that this game is soccer.

The arguments in favor of such a revision come mainly from the ontological perspective: we claim that, with our modified parametrical CRs theory, you can avoid the rigidity of social practices and the explosion of types, i.e. the addition of one player does not entail – if within the parametrical range – the automatic creation of a new type of soccer. On the other hand, this approach, by forcing you to investigate dynamic games played for real (i.e. tokens of games, real life matches) brings into account more methodological and epistemological complexities.

Another example is the following: consider the rule that constitutes the size of a tennis court: in case you have a rule saying that the surface of the tennis court is the same as the universe you are not (only) constituting another game – universal tennis – but you are going into a new game by entering the wrong value into one of the parameters of a CR (i.e. the CR that rules the tennis court size).

3. New Horizons: Violating Constitutive Rules Our new parametrical framework enabled us both to modify a game without exiting the type by way of changing the value of the parameter according to the range and to allow for violations of CRs (i.e. for breakable CRs). The theoretical move is to say that violations occur when the value that we choose does not fall out the range of the parameter itself but is different from that value of the parameter that is specified by the codified constitutive rule.

This feature of the theory enables us not only to allow for CRs’ violations but also to classify the possible ways of exiting a game. From the standpoint of view of the gaming practices, we will distinguish a semantic exit vs. a practical exit.

3.1. Ludic Point of View: Semantic Exit vs. Practical Exit Consider again the rule concerning the numbers of the players in soccer in its parametrical form:

(*N° CR) soccer is played X vs. X.

Now consider the following examples:

(1*) soccer is played 1 vs. 1;

(2*) soccer is played 100 vs. 100;

and compare them with what, in our theory, is a violation of the codified soccer rule concerning the number of the players (11 vs. 11), e.g.:

(3*) soccer is played 12 vs. 12;

(1*) is a *semantic exit* from the game of soccer because it goes against the definition of soccer played as a *team sport*⁹.

(2*) is a *practical exit*. Suppose we manage to play something like a soccer game: all the FIFA official rules apply except for the fact that we play in 100 vs. 100. The game will be much different from the FIFA soccer as we know it, dribbling will become less useful and the whole game will be much more “crowded”, a sort of scrum in rugby. Playing the game in 100 vs. 100 is not impossible nor it goes against the definition of soccer as a sport team, but it will create many physical and practical problems in order for the game to go on as we expect it to go. We label such a way of stretching too far a parameter, *practical exit*. Practical exits lead to a new and interesting question: what are the right parameters for soccer (and the other games in general)? How do we recognize them¹⁰?

(3*) is, as we know already, a violation of the codified 11 vs. 11 soccer rule concerning the number of the players. Nonetheless, it is not enough to exiting the game. All violations, in fact, are nothing but variations of the codified value of the parameter, but they do not fall out of the value assignment for it.

Comparing how the different theories deal with inviolability, we can say that there is a violation when - in our parametrical theory - the value assigned to a parameter departs from the value that corresponds to the codified version or the most played version (in soccer 11 vs. 11).

In that way our new theory admits as a well-formed rule what was an impossible violation in the previous theory. Consider 7 vs. 7 soccer: this was not a violation in the standard theory, it was not even soccer but another new type of game with no connections to 11 vs. 11 soccer. In our parametrical theory 7 vs. 7 is soccer and, given the historical perspective our methodology opens to us, we can see it as a violation of the original rule of 11 vs. 11 that is now included in soccer.

We agree with the standard theory that, in some cases, we go too far and exit a game. With this regard, our theory distinguishes a semantic exit from a pragmatic exit.

We showed that a framework in which CRs are breakable and we provided examples for it. We also addressed the new topic of exiting a game. Now it is time to consider some criticisms to our account.

One of the possible objections that can be raised to our approach is that we did not provide examples in the domain of chess that, in the history of philosophy, is somehow the paradigmatic example of practice constituted by rules and we relied instead only on soccer and tennis¹¹.

⁹ You may still say that 1 vs. 1 can be training to play soccer and so is 1 vs. 0 (say when you practice shooting corner kicks). Nonetheless, in order to say this, a theory to explain training as a subgame is needed.

¹⁰ This is a rather complicated question deserving a whole methodological discussion. Here we will give just a short answer regarding soccer as a team game. The least number is 3 because even in 2 vs. 2 there are chances for the game not to be the real team game of soccer: the team will have a goalkeeper and a field player, but practically this amounts to 1 vs. 1 shooting towards a keeper (we may say it is 1 +1 vs. 1 + 1). So we need at least 3 players per team so that the two players remaining - not counting the keeper - are able to play as a team. The upper bound is set to 15 by way of looking at the number of players that other similar sports played on similar fields have: Gaelic soccer, rugby, hurling are all played 15 vs. 15 all involves having some sort of ball crossing a line and are played on a field with approximately the same length as the soccer pitch.

¹¹ To be more accurate, chess are paradigmatic in the prehistory of the concept of constitutive rules and in the legal philosophy development. Searle uses pretty often examples from American football such as the touchdown or the realm of institutions (creating money, being a president, etc).

3.2. Comparing Theories

4. Conclusion: Vagueness in the Parametrical Constitutive Rule Theory

We claim that this “argument” clearly misses the target. Chess are a sort of naturally rigid practice (they can be perfectly played in absence of any physical structure - i.e. blind chess) so, if you conceive all the gaming practices as chess-like practices, CRs will hardly have a chance of being violable.

For the chess maniacs we can, nonetheless, provide two (thought experiment) examples: the first will show that *there are* parameters even in chess; the second will show that, not only there are parameters in chess, but that there are some value assignments that count as wrong assignments.

Imagine you are playing chess and that you are asked to play with a 7x7 chessboard or a 10x10 chessboard. This is enough to show that there can be parameters in the CRs of a game (the default value for the number of squares is 8x8).

Now reflect on your reactions to this proposal: you may regret 7 or 10 as a correct the value for the parameter: this proves that there are parameters in chess and that some wrong assignment that may lead to a violation or even to exiting the game.

To reinforce the point that even in chess it is possible to have an assignment in the parameters that counts as a violation of a rule (and so, according to our theory, which does not entail exiting the game), consider a game of chess where both the players invert the parameter of the starting position of the king and the queen. King and queen have a starting position parameter that in official chess are – from the white’s player viewpoint – 1D (Queen) and 1E (King). In this new version you play chess with 1D-King and 1E-Queen. Despite the fact that you are switching parameters, it is still a game of chess where all chess moves are possible and chess strategies are the same. You may need to adjust some of them to the new starting position (think about castling) but it is not the case that some moves are going to be impossible¹².

With the new parametrical theory of constitutive rule we are proposing, problems of vagueness will obviously arise as a direct consequence of having replaced a precise value with a parameter to which we assign a range of values. Of course it is possible to debate on the range of values tolerated by the parameters before you fall out of the type of a game and it is worth investigating the rules of a game which are more affected by vagueness. The very possibility of discussing these issues proves that our work is going in the right direction¹³.

¹² Compare this to playing something similar to chess where you switch the king with the pawn in front of it. Here you will lose any chance to castle.

¹³ We would like to thank the audiences of the “Workshop Society and Values” and “SIFA 2012” conferences as well as the anonymous referee of this journal for helpful comments and objections to our paper.

REFERENCES

- Benoist, J. (2003), "Quand finissons-nous un jeu?", in P. Di Lucia (editor), *Ontologia sociale, potere deontico e regole costitutive*, Quodlibet, Macerata, 2003, pp. 191-196;
- Conte, A. G. (1983), "Paradigmi d'analisi della regola in Wittgenstein", in Conte (1995), pp. 267-312;
- Conte, A. G. (1989), *Filosofia del linguaggio normativo I*, Giappichelli, Torino, 1989;
- Conte, A. G. (1995), *Filosofia del linguaggio normativo II*, Giappichelli, Torino, 1995;
- Conte, A. G. (2000), "Nomotropismo: agire in funzione di regole", in Conte (2001), pp. 893-920;
- Conte, A. G. (2001), *Filosofia del linguaggio normativo III*, Giappichelli, Torino, 2001;
- Guastini, Riccardo (1983); "Teorie delle regole costitutive", *Rivista internazionale di Filosofia del diritto*, 60, 1983, pp. 548-564;
- Rawls, J. (1955), "Two Concepts of Rules", *Philosophical Review*, 64, pp. 3-32;
- Ross, A. N. C. (1968), *Directives and Norms*, Routledge & Kegan Paul, London, 1968;
- Roversi, C. (2007), *Pragmatica delle regole costitutive*, Gedit, Bologna, 2007;
- Searle, J. R. (1969), *Speech Acts. An essay in the Philosophy of Language*, Cambridge University Press, London, 1969;
- Snyder, A. (1971), "Rules of Language", *Mind*, 80, 18, 1971, pp. 161-178;
- Żelaniec, W. (2003), "On the very idea of constitutive rule", trad. it. di Edoardo Fittipaldi, "Sull'idea stessa di regola costitutiva", in P. Di Lucia (editor), *Ontologia sociale, potere deontico e regole costitutive*, Quodlibet, Macerata, 2003 pp. 155-79.

EMANUELE BOTTAZZI

ROBERTA FERRARIO

Istituto di Scienze e Tecnologie della Cognizione di Trento

emanuele.bottazzi@loa.istc.cnr.it

roberta.ferrario@loa.istc.cnr.it

APPEARANCE COUNTING AS REALITY? SOME CONSIDERATIONS ON STABILITY AND UNPREDICTABILITY IN SOCIAL INSTITUTIONS

abstract

Social institutions are considered as generally stable entities but, at the same time, prone to unpredictable and even dramatical changes. The paper draws some considerations on this issue by means of an analysis of the notions of validity and objectivity, seen in the light of the critical situations that institutions may face. Consider the case of an irregular election that leads anyhow to a result accepted by authorities and population at large. Such an election seems to be valid and not valid at the same time. These contrasting intuitions reflect the twofold nature of institutional reality: it depends on one side on what is accepted by those who are dealing with it, but on the other side also on what it is actually going on according to its rules. The article frames this problem in terms of response-dependence. In this respect, it follows, but also expands, Hindriks (2006), to reach the conclusion that institutions are easily prone to fall apart. Anyway, not in all cases this actually happens; on the contrary, institutions seem to be quite stable. We propose that what makes them stable, but also unpredictable, is, in a sense, a sort of misunderstanding among the involved agents. Explaining such misunderstanding amounts to analyze the notion of objectivity in social reality. We argue that objectivity is more a requirement than a feature of institutional systems, being institutional objects mind and context-dependent.

keywords

Validity, objectivity, crisis in institutions, stability, unpredictability

Suppose a significant number of voters is voting in the wrong way. For example, they make marks on the ballots that compromise their secrecy. According to the rules, their vote is null. But, in order to be considered as null by the institution, these votes must be recognized as null by some other “relevant” people (poll watchers, judges of some court). Suppose that, for some reason, none of them recognize the violation and suppose, on top of this, that none of the voters realize that her/his vote was wrongly expressed by the poll; what we have here is a situation in which, from a god’s eye perspective, the entire election is null, but from the perspective of the considered group the election is valid. Valid, exactly because it appears to be valid.

The situation is somehow surreal, but it makes a point: institutional reality is twofold, but in a very peculiar way. There is a significant difference between what is really going on according to the rules and what is believed and accepted by those who are dealing with the institution. Nonetheless, what is believed and accepted is crucial in assessing what is really going on in the institution. As a further complication, we cannot avoid to use some strong sense of objectivity: in many situations we feel that it is important to know what actually happened according to the rules, as we want this to have an influence on our decisions.

What we would like to claim is that in the domain of social ontology somehow both views are not only needed but somehow inevitable to retain. If we lost the former, we would have to say that what is objective has no role in what is ontological in social reality, and then no subsequent discovery of error can have any role in establishing who won the elections, if we lost the latter we would end up saying that what is subjective is not relevant, undermining the very sense of social activities, including polls. Then, what is needed in order for such kinds of situations to be “fully” valid?

This seems to be a problem that is peculiar of the social domain, since apparently it does not apply, at least not in the same way, to contexts such as the natural one, or, to be more precise, to the so-called *mesoscopic level*: if we wrongly believe that “the cat is on the mat”, such belief has few effects on the actual position of the cat. On top of this, the example above seems to clearly belong to a subdomain of the social, that is the institutional one. Institutions, as many have pointed out, depend, in order to

exist, on rules that define some of their own relevant aspects; e.g., voting rules are essential for polls.

One of the main supporters of the latter view is certainly John Searle (1969, 1995, 2010); take as an instance his famous analysis of constitutive rules, that has generated an important field of studies. But one of the criticisms that has often been moved to Searle's account is that it is too centered on positive situations in which agents quickly agree and are successful in their cooperation, but is less effective when analyzing critical cases, cases as this one, that may be puzzling, but also revealing. To be completely honest, such criticism was only supported by a minority¹ of scholars and, in general, even if it seems obvious that we often face critical situations in our social and institutional activities, there is not much work carried out in analytic social ontology about this very fact.

Taking care of this “dark side” of social ontology is crucial, also because rules themselves may constitute a threat to the understanding of the ontology of the world they create. Aside of their misapplication, as in our poll case, contradictions or faults may be part of the rules themselves and this makes it difficult to assess what is valid in some specific institutional domain. One of the desired outcomes of this analysis is to establish some pillars in order to draw a link between these two dimensions (misapplication and faults in the rules), as these are important assets for the understanding of the dynamics of institutions. Our starting point will be the notion of acceptance-dependence as developed by Frank Hindriks, and especially the version he proposes of the concept *constitutional*. This analysis, complemented with our considerations about the “dark side”, will allow us to show some of those traits that make institutions so stable and, at the same time, so prone to sudden, unpredictable changes.

There are many possible ways of conceptualizing the problem of acceptance in social ontology. An interesting and somehow new way to give a stronger grasp of the issue is by re-describing it in terms of the recently and widely debated notion of *response dependence*². The term “response-dependent concept” was introduced by Mark Johnston, (1989, 146, footnote 8) and, in a nutshell, is aimed to capture the idea that there are concepts that are *intrinsically*, or *a-priori*, determined by our responses; they “are tailor-made for creatures like us who are capable, as many intelligences may not be, of certain responses: capable of finding things smooth to the touch, bland to the taste, red to the eye”³. The classical example are colors⁴; something is **red** if and only if, in normal conditions, this very something is disposed to look this way to *normal* subjects. It is important to stress that the thing has to look red to a normal subject in normal conditions, since the idea is to have some kind of objectivity, supposing that all these conditions can be specified in a germane manner; that is, it is to select the conditions in a way “to ensure that no other circumstances could have given the judgment formed a greater credibility”⁵.

Frank Hidriks (2006) was perhaps among the first scholars who attempted to make a systematic link between sociality and response-dependence, by introducing the notion of *acceptance-dependence*. To this aim, he starts by considering Crispin Wright's (1988, 1992) specific approach on the issue (called *judgement-dependence*), since he sees the latter as more similar to acceptance-dependence than any of the other available accounts. The main difference between the two approaches is that *judgement-dependent* concepts depend on counterfactual responses of individuals while, in contrast, *acceptance-dependent* concepts depend on the *actual* responses of *groups* of people. The counterfactual idea of judgment-dependence amounts to saying that the judgments of standard subjects under standard

1. Acceptance Dependence (and its Dark Side)

¹ For example (Balzer 2002) and (Schmid 2009).

² There are many theories around on this notion, for recent reviews of the literature see (Yates 2008) and (Gundersen 2010).

³ (Pettit 1991, 587).

⁴ Even if it has been argued that colors may be not response dependent, e.g. in (Johnston 1989), we will use them here just as illustration of the notion.

⁵ (Wright 1992, 109).

conditions are destined to be correct: something is *red* if and only if an idealized subject has the red-judgment in standard situations. Instead, if we take a social concept, for instance *popular*, we have to consider the *actual* responses by groups of people. In order to assess if the concept of *popular* applies to Ada, we have to see if Ada is *actually* liked by a relevant number of people, rather than trying to find out whether an idealized subject has the popular-judgment in standard circumstances. Another relevant difference is the *contextual character* of the notion of acceptance: being popular is something generally true with respect to some group of people, but false if referred to another.

Hindriks provides an interesting analysis of the concept *constitutional*, that we want to use as a scaffolding for the analysis of validity, as we mentioned in the introduction. But being *valid* in our sense is not necessarily being “legally valid in a specific legislation”; instead, we want our analysis to be more general, we want it to encompass what is decided as the final outcome of an election, even if we would like this outcome also to be taken into account in the analysis. The analysis of *constitutional* is therefore useful but, at the same time, it can not be sufficient to characterize our notion, since, at least as it is presented in Hindriks’ account, it seems to mean something very similar to “legally valid in a specific legislation”:

Hindriks’ account moves from a criticism of Johnston’s:

Johnston claims that constitutional is judgment-dependent and offers the following analysis:

“Thus the concept of a US state or federal law’s being constitutional is... the concept of the Supreme Court’s not being disposed to ultimately regard it as unconstitutional. (Being constitutional is the default condition)” (1993, p. 104; emphasis in original). [...] Johnston’s analysis is problematic because it is compatibility - or rather, incompatibility - with the constitution that matters, instead of what the Supreme Court is disposed to do. (Hindriks 2006, 486)

As Hindriks remarks later, “constitutionality is fixed independently of what the Supreme Court is disposed to do” (Hindriks 2006, 487), it is what is written in the Constitution that marks the difference between what is and what is not constitutional. In this case, what makes something constitutional is anyway non-counterfactual, being constitutional something related to the citizens contextual and *actual* acceptance of the Constitution. *Actually* accepting the Constitution is, indirectly, providing *actual* criteria of constitutionality. Said that, he highlights another important point. There are problematic cases, where it is difficult to assess whether something is constitutional or not. In these cases the authority of the Supreme Court is what determines whether or not a law is constitutional. In these latter cases, anyway, actual rather than counterfactual responses are crucial for the application of the concept constitutional.

Now we move to consider Hindriks’ definition of acceptance-dependence. In his paper he provides slightly different definitions of it, but the definition we are interested in is the one that is best suited for the concept constitutional:

[AD*] x is F in context C in context C group G accepts a rule R or authority T has declared that a rule R is in force and according to this rule x is F . (Hindriks 2006, 492).

For Hindriks the formulation above “explicitates the relation between rules, our actual responses, and concepts that depend on them in an indirect manner” (Hindriks 2006, 492), that is to say that it can be applied to the concept of constitutional, because it covers both the unproblematic cases with sets of

accepted rules, when it is sufficient to refer to the Constitution (expressed in the first disjunct of the right part of the bi-conditional), and the problematic ones, when we have to recur to the authority of the Supreme Court (the second disjunct of the right part of the bi-conditional).

We believe that even if this analysis has its own merits of clarity and it is quite intuitive, it is not applicable to our case, namely to define the concept *valid*⁶ that we would like to use in our analysis. The reason for such inapplicability is in the “or” in the right part of the bi-conditional. Let’s consider the poll example provided at the beginning of the paper. If Hindriks’ definition is correct, then we would not feel any problem in accepting the validity of the electoral outcome. The outcome is perhaps constitutional, but it is not “fully” valid, since what *actually* happened is against what has been fixed in the rules. In our intuition, in order to have “full” validity, what is fixed in the agreed rule and what comes as the final decision of the authority have to be in accordance, something like: *x is valid in context C in context C group G accepts a rule R and authority T has declared that a rule R is in force and according to this rule x is valid.*

This is to say that valid is a response-dependent concept with a “stricter” definition than the one of constitutional, since the disjunctive formulation implies the conjunctive: it holds, in principle, in less cases. What are the consequences of this line of reasoning? Validity for an institution is the very heart of the institution itself. An institution that is perceived as not valid is bound to failure, since the mismatch between what the rule, if significant, says and how it is applied, makes the whole system unable to coordinate people’s action and then no more acceptable for them. If, in the example above, everyone involved in the poll woke up and realized that the outcome of the election, established by authority, did not reflect what actually happened, then there are good chances that the whole system would fall apart. This is, we believe, something inherent to every institutional system and also, so to speak, one of its most problematic “dark sides”⁷.

This dichotomy between rule and authority manifests itself very often in the life of an institutional system, but anyway not in all cases everything falls apart. On the contrary, institutions seem to be quite stable⁸. We believe that what makes them stable, to put it in a provocative way, is *misunderstanding*.

What we have established so far is that there is a certain requirement of objectivity of rules and their application. This is to say that, for some institutional entity to be considered as valid in the full sense, rules have to be correctly applied in accordance with what they say. Then, we suggested that breaking this link between rules and application is potentially harmful for the institutional system as a whole: if members of the community perceive that the system is not valid anymore, they cease to accept it and this will very likely make it fall apart, or will force a substantial change. Let us elaborate a bit on this to conclude, from the following considerations, that it is exactly because agents involved do not have a clear idea of what is going on in their institutions that institutions are generally stable and, at the same time, prone to unpredictably end up in critical situations.

Objectivity in our view is more a requirement than something that is effectively and always present in our institutional world. But before going into that, we have to illustrate why objectivity is important. The main reason is that what is believed can have a crucial effect on what is accepted. Acceptance

6 Some of the criticisms that we will provide for valid hold also for constitutional, but we will not concentrate on them in this work, due to space limitations, but also because constitutional is not the focus of the paper, for reasons expressed earlier in this section.

7 There are other elements in this dark side; for example rules can be contradictory, as in the case of the Discursive Dilemma (List and Pettit 2004), but we will not deal with such elements in this paper.

8 This point is quite accepted in literature, from the first studies in sociology (Weber 1968) to the latest in philosophy of sociality (Searle 2010).

2.
Lack of
Objectivity
Makes
Institutional
Systems
Stable, but
Unpredictable

and belief are often considered, by philosophers, as distinct⁹. Normally, you can cheat in a game only if I don't know that you are cheating. This is not entirely correct, since I can go on playing even if I believe that you are cheating. Among the various points made to support the distinction between belief and acceptance, (Cohen 1992) says something interesting: belief and acceptance differ in that acceptance arises from adopting a policy to achieve a particular goal. I accept that p if I have the policy of postulating that p. In our game, I can accept your cheating because I will gain some advantage by doing it, for example, I simply can be happy to let you win. If I accept your cheating, it means that what we both accept is the fact that we are playing regularly, but each one of us believe that the game is not valid, or, better, that it is not *fully* valid, since the cheating has not been declared¹⁰ and accepted yet. But if I discover the cheat I *can* halt the game, in the sense that I have good arguments to make you accept, for example, that the game is null or that you ought to be sanctioned.

There are other motivations that can force agents to declare the misapplication and try to convince the others to accept it and, then, to act accordingly. Aside malicious interests, one of the main motivations is the understanding of the fact that, in some cases, the misapplication could be harmful for coordination. For example, I can realize that my opponent is misapplying the rules in such a way that I do not know how to reply with another move. If I feel that I am in a deadlock and I cannot go on playing, we can say that I feel forced to declare the misapplication.

This is enough for arguing that objectivity, as we intended it here, plays a crucial role in institutional systems as a requirement. But this does not mean that this is also a necessary condition for an institutional system to be an institutional system, this just tells us that agents *can* use this requirement in order to make other agents accept that some misapplication has taken place and change their behavior accordingly.

For some theorists we have a sort of objectivity in our knowledge of social reality. The debate is very complex¹¹, and we will suggest just two quick lines of attack to this position. The first line deals with the relationship between the ontological and the epistemological dimension.

According to Searle, the institutional reality is *ontologically subjective*, i.e. it depends, for its existence, on agents' mind. On a par with this thesis, we have the other one, namely that the institutional reality is also *epistemically objective*. This means that "it is, for example, an epistemically objective feature of this thing that it is a screwdriver; but that feature exists only relative to observers and users, and so the feature is ontologically subjective." (Searle 1995, 10).

We believe that this objectivity is not as firm as Searle would want it to be. The dependence of social entities on agents' minds, rules and their application makes things less objective than one could think and leaves room for misunderstandings or lack of proper knowledge of what is going on in the institutional system. To see it, let's go back to the comparison Searle makes with artifacts on one side, like screwdrivers, and social entities on the other. Suppose we accept that there is something like "an objective feature" that allows us to classify this thing as a screwdriver. Nonetheless, screwdrivers can be used in different manners, not even imagined by the users and the creators of

⁹ For example (Cohen 1992), (Bratman 1999) and (Tuomela 2000).

¹⁰ The underlying idea is that collective acceptance is carried out by means of declaratives. For some recent work on declaratives see (Searle 2010) and (Tuomela 2011).

¹¹ See (Searle 1995, 2010), but (Thomasson 2003) has even a more extreme position on that. (Guala 2010) labels this position as infallibilism, meaning that, being social kinds a product of stipulation, then at least some (relevant) members of the community that makes such stipulation cannot be wrong about them: "we have certain forms of epistemic privilege with regard to our own institutional [...] kinds, protecting us from certain possibilities of ignorance and error (Thomasson 2003, 580)."

the screwdriver. This could happen in such a way as to make it difficult to recognize this thing as a screwdriver: it is so out of context that we cannot understand anymore what it is. This holds, in principle, for all the things we were willing to classify as screwdrivers, but, actually, these different uses pose a threat neither to the new nor to the old context. There is just a new concept for classifying things with the same old shape. This boils down to say that I can use the very same thing as a screwdriver and as a nail. But to deal with the institutional world is a different matter. Institutional entities have rules that concur to define them, we can say that the former have a content that establishes their use, since they are themselves made up of rules, not out of physical matter as screwdrivers. On top of this, they are related to a context of other rules and such context is much stricter. If we go out of the original context, there is always the possibility that we cannot understand what is the new application anymore, as we also suggested when talking about objectivity as a requirement and its role in coordination.

The second line of attack is that acceptance and belief have to be considered as spreading all over an institution, especially through the agents involved, who are, in the end, those that put social reality into existence. As we said for the chess example, each “move” of each agent involved in the institution has to be accepted by the other agents. But this means that in order to assess what is valid in the full sense and what is not, we should check every agent’s action, epistemic attitude and act of acceptance. Moreover, institutions as states have a huge amount of rules and dispositions and they perform, via their relevant members, a vast amount of institutional actions. It is then very likely to imagine that each agent has a limited amount of knowledge on what is going on in the institution.

How do members of an institution cope with this state of ignorance or misunderstanding? There are many possibilities that need to be explored, a tentative solution could be that they rely on trust. They somehow trust each other and the authority for the application and control of the application of the rules and for resolving difficult matters. But also authorities (as *relevant members*) have a limited, even if, maybe, greater, amount of knowledge of what is going on in the system, and they are also forced to trust the system. The fact that we trust that, at the end of the day, rules are more or less applied in the correct way, exactly because we don’t have a precise way to establish it, is what makes institutions stable. But this can enhance the possibility, for some of the previously mentioned reasons, of a serious divorce between rules and their application. Normally, this divorce has no dramatic effects on the institutional reality, as it is not even recognized by its members and agents tend to rely more on the application than on the rules themselves. We said also that the system is unpredictable for the same reasons. This is to remark how unlikely it is that the agents realize the effective status of the system, in this case that there is a problem in the rule-application axis. And also when they realize it, it could be that they have no interest in making this explicit or accepted. Even when explicitly accepted, they can be happy with it, meaning that it can have no interference with their policy. It is only when there is a convergence of this very complex chain of beliefs and acceptances with some specific crucial interests, like the need for coordination, that the system falls apart.

REFERENCES

- Balzer, W. (2002), "Searle on Social Institutions", *Dialectica*, 56 (3), 2002, pp. 195–211;
- Bratman, M. (1999), *Faces of Intention*, Cambridge, CUP, 1999;
- Cohen, L. (1992), *An Essay on Belief and Acceptance*, Oxford, OUP, 1992;
- Gundersen, E. B. (2010), "Dispositions and Response-Dependence Theories", in R. Poli, M. Healy, and A. Kameas, (Eds.), *Theory and Applications of Ontology*, (vol. 1, pp. 119–134), Springer, 2010;
- Hindriks, F. A. (2006), "Acceptance-Dependence: A Social Kind of Response-Dependence", *Pacific Philosophical Quarterly*, 87, 2006, pp. 481–498;
- Johnston, M. (1989), "Dispositional Theories of Value", *Proceedings of the Aristotelian Society*, 63, 1999, pp. 139–174;
- Johnston, M. (1993), "Objectivity Refigured: Pragmatism without Verificationism," in J. Haldane and C. Wright (eds.) *Reality, Representation, and Projection*. Oxford, OUP, pp. 85–130, 1993;
- List, C. and Pettit, P. (2004), "Aggregating Sets of Judgments: Two Impossibility Results Compared", *Synthese* 140, 2004, pp. 207–235;
- Schmid, H. B. (2009), "On Not Doing One's Part", in *Plural Action: Essays in Philosophy and Social Science*, pp. 47–58, Springer, 2009;
- Searle, J. R. (1969), *Speech Acts, an Essay in the Philosophy of Language*. Cambridge, OUP, 1969;
- Searle, J. R. (1990), "Collective Intentions and Actions", in P. R. Cohen, J. Morgan & M. E. Pollak (Eds.), *Intentions in Communication* (pp. 401–415), MIT Press, 1990;
- Searle, J. R. (1995), *The Construction of Social Reality*, New York, The Free Press, 1995;
- Searle, J. R. (2010), *Making the Social World: The Structure of Human Civilization*, Oxford, OUP, 2010.
- Thomasson, A. (2003), "Realism and Human Kinds", *Philosophy and Phenomenological Research*, 68, 2003, pp. 580–609;
- Tuomela, R. (2000), "Belief Versus Acceptance", *Philosophical Explorations*, 3 (2), 2000, pp.122–137;
- Tuomela, R. (2011), "Searle's New Construction of Social Reality", *Analysis* 71 (4), 2011, pp. 706–719;
- Pettit, P. (1991), "Realism and Response-Dependence", *Mind*, 100(4), 1991, pp. 587–626;
- Weber, M. (1968), *Economy and Society*, Bedminster Press, 1968;
- Wright, C. (1988), "Moral Values, Projection and Secondary Qualities", *Proceedings of the Aristotelian Society* 62, 1988, pp. 1–26;
- Wright, C. (1992), *Truth and Objectivity*, Cambridge, Mass, HUP, 1992;
- Yates, D. (2008), "Recent Work on Response-Dependence", *Philosophical Books*, 49(4), 2008, pp. 344–354.

GAETANO ALBERGO

Università degli Studi di Catania

gaetanoalbergo@yahoo.it

DOES ONTOGENESIS OF SOCIAL ONTOLOGY START WITH PRETENCE?

abstract

Rakoczy and Tomasello follow Searle in claiming that rule games need status function assignment and constitutive rules. But, in the case of pretend play, it is not easy to put together these notions with the natural world knowledge necessary to engage in it. If we consider the pretended scenario as a possible world, metaphysically possible, then, how can we abandon the natural necessity implicit in it? The rules of pretend-inference can have a robustly objective status. On this view pretence stands to pretending as truth stands to belief.

keywords

Status function, pretence, non-conceptual imagining

Gelman and Byrnes told us, in Chomsky's terms, that "[w]e can determine how languages and conceptual systems are constrained by examining the forms and meanings that children construct, and which errors they fail to make" (Gelman and Byrnes 1991, 3). Thus Billman suggested that we should compare and test psychological models of structure and processing of concepts by examining the function from "learning instances plus the target items to categorize" to "the set of possible category judgments" (Billman 1992, 415). Then the actual extension of the category is taken to be determined by how the learner is inclined to classify new examples. But I believe that experiments need to be designed and interpreted with it in mind that the cognitive systems are designed by evolution and tuned by experience to find real-world substances, not random logically possible ones. I think *pretence* is a special case of conceptual tracking. Pretending is changing the nature of perceptual inputs at will. I suggest the exercise of this kind of agency can enable subject to experience the "refractoriness" of reality, experience the constraints that the reality sets on what they can experience.

According to Hannes Rakoczy and Michael Tomasello the ontogeny of social ontology starts with pretend play. This is the voluntary transformation of the here and now, the you and me, and the this or that, along with any potential action that these components of a situation might have. Thus, *pretence* presents a cradle for children's development into institutional life, but, in order not to be confused by a parent's pretend acts, the child must interpret *pretence* events as what they stand for. According to Rakoczy and Tomasello (R&T) this is so because on embryonic form these phenomena exhibit the logical structure of the conventional creation of institutional facts. Following Searle, they recognize that, in contrast to brute facts "out there", *institutional facts* hold only by virtue of a social, conventional practice and have the logical form "x counts as y in a context C". Every stipulations of the *pretence* scenario involve a *status function* assignment. So children do not get confused about reality because of the symbolic nature of status function. Quoting Searle "The sense in which symbolization in [a] broad linguistic sense is essential to all institutional facts is that the move from x to y in the formula 'x counts as y in C' is already a symbolic move" (Searle 1999, 155).

I agree with R&T that *pretence*, if it is an institutional fact or not, involve normative aspects, that is, what is appropriate, what is a mistake, or highly inappropriate, in a given context. For example, Rakoczy has shown that children from around 2 years old not only are proficient at acting according to jointly set-up fictional stipulations in the context of shared pretend play, but when a third party confused pretence identities and thus made mistakes, children leveled protest and critique (see Rakoczy 2008). But we don't have a conclusive evidence that early *pretence* involves the symbolic function. Some evidence stems from old studies showing correlations between early acquisition of words—symbols for referents— and use of pretence gestures (e.g., Bates, Benigni, Bretherton, Camaioni, and Volterra 1979). On the contrary, in a recent research, Angeline Lillard claims that young children's understanding of *pretence* is heavily guided by their ability to read gestures out of context, perhaps only in the face of supportive social signals indicating *pretence* rather than by employing a symbolic function (Lillard 2001). I don't think that symbolic function is always a necessary condition for an institutional fact. In fact, we can give an independent account of normativity and not-contradiction in *pretence*, without the symbolic function.

If pretending is a case of "thinking of *x* as *y*", *y* is the way *x* is described, but, if we follow Roger Scruton's account of imagination, broadly conceived, then *pretence* involves thinking of these descriptions as appropriate in some way to the primary object (see Scruton 1974). Pretending may be compared with uttering a sentence, as distinct from asserting it. We know that not everything that is said is also asserted. I may practice my pronunciation and simply utter a sentence, so I represent a state of affairs, but I don't thereby assert that the state of affairs is real. I merely represent a possibility without asserting it, and we know, as empirical evidence tells us, that children don't believe what they pretend. Nor do I hesitantly assert what the sentence says. Pretending cannot be analyzed in terms of belief. The content of my belief can be expressed by a sentence, in this case sentences are being used to say something. In a secondary case, as in elocution, sentences are being treated more as patterns of sound than as verbal symbols. If pretending is a case of "thinking of *x* as *y*", then to think of *x* as *y* at least involves the entertaining of the proposition "*x* is *y*". So, if someone pretends that a banana is a telephone, then he has a disposition to say "I think of the banana as a telephone". But, it is not sufficient for the truth of his pronouncement that he should have a disposition to entertain this proposition if he always immediately rejects it as false. Actually, I may think of *x* as *y* while knowing that it is untrue. On the other hand, with symbols we are limited to the dimension of falsity. By definition it's always false that *x* is *y* if the first is a symbol for the second. And, clearly, marking the pretend representation as false does not even begin to capture the semantics of early *pretence*. In a "mark as false" account a child infers that "mummy pretends that (this banana is a telephone)*false*". Then she gives him the banana and says "the telephone is ringing". The child represents this as "mummy pretends that (the telephone is ringing)*false*". But, which telephone? This problem does not arise if we merely entertain the proposition "*x* is *y*". The failure of reference in the speech-act is a consequence of my entertaining unasserted the existential proposition whose truth is a necessary condition for the truth of "this banana is a telephone". Of course, we know that the pretence scenario is not true, but, on the other hand, there are precise true-conditions to decide if pretending *P* is a case of *pretence*. In my opinion, to say that *P* is *pretence* is to make a normative rather than a descriptive claim, and children are sensitive to what is pretence. In this way I reject the contrary view that a child can understand what is *pretence* only by understanding that someone is pretending something. *Pretence*, I think, is a rational activity, then, what is to judge an unasserted description to be appropriate to a certain object? Pretending is a way of going beyond the "given", the primary object, and producing descriptions which one is unprepared to assert. But this is not sufficient. It is necessary that the description be entertained because of his appropriateness.

It is true that *pretence* does involve treating objects as if they were something else, and R&T quote Searle “In the limiting case, we can use the object itself to represent the y status function” (Searle 1999, 155), but this cognitive activity is more sophisticated. For example, following Alan Leslie (1987) we can recognize three basic forms of *pretence*, with their semantic properties:

1. Object substitution (referential opacity)
2. Attribution of pretend properties (nontailment of truth-falsehood)
3. Imaginary object *pretence* (nontailment of existence)

It has been argued that pretending 2-year-olds understand four features of *pretense*: pretend stipulations, causal powers, the suspension of objective truth, and an unfolding causal chain. The current situation might contain a toy horse or an empty cup. Then, for example, memory systems are addressed, returning information on entities that are perceptually similar (e.g., on horses) or on the functional properties of the object (e.g., on containing). This leads to *pretence* based on perceptual similarity or on functional connection.

Markman and Abelev (2006) suggest that pretence is not unique in permitting 2 and 3-year-olds to recognize multiple object identities. In their study children were as good at recognizing unusual functional use as they were pretence, while still failing the appearance-reality task where deceptive stimuli were used across tasks. Their interpretation is that children understand multiple object identities better when an object’s intrinsic identity is contrasted with its relational role, that is an extrinsic *object properties*. So, children are able to distinguish extrinsic object properties from intrinsic ones (function vs. category-membership) better than they can distinguish superficial object properties from deep ones (appearance vs. category-membership).

If we follow this account, then, conceptual factors, such as the intrinsic or extrinsic nature of the alternative identity, help shape children’s performance on multiple object identity tasks.

Now we can try to answer the question Scruton posed about appropriateness of descriptions.

We have two options:

- (i) Keil-Leslie: Anchoring is straightforward for immediate pretense; it proceeds by means of a best formal match between expressions in the pretend and current perceptual representations
- (ii) Searle: The status and its accompanying function go beyond the sheer brute physical functions that can be assigned to physical objects

So, we stay between Scylla of highly structured scenario and Charybdis of an umbrella-term that covers many different phenomena. We need to define our domain of search a bit more precisely.

In these situations we are working on the identity of x and y , and, my claim is, if we consider the new scenario as a possible world, metaphysically possible, then, how can we abandon the *natural* necessity implicit both in the stipulation and in successive possible implications¹? Children’s understanding of the causal consequences of a pretend action is a routine by-product of a simple processing rule: to understand the consequences of a pretend action, assume that the entities or substances whose

¹ I follow Kit Fine in arguing that thinking about identity is a necessary condition to talk about metaphysical modalities, and this cannot reject physical notions. Rather than to give up the idea that there are any natural necessities, thinking about identity, I believe, should lead us to adopt a more discriminating view as to which natural necessities are metaphysical contingent.

existence is stipulated are subject to the same causal principles as their real-world equivalents (see Paul Harris 2000; Alison Gopnik 2009). Normativity is constitutive of the scenario, but it is strongly appreciated in its possible consequences. This shows that children need to draw on familiar, real world causal principles. This evidence is not compatible with the way Searle uses the notion of *observer-dependent fact*. According to Searle the symbolic nature of a similar fact is not a product of physical virtues. Moreover, if we follow R&T in claiming that *pretence* is probably the ontogenetically primary factor for institutional life, then, should we adopt such a conclusion as: *pretence* is “a matter of status functions imposed according to constitutive rules”? (Searle 2005, 9) But, which constitutive rules? Even R&T recognize that in this case “the “y”, is not, as in the case of rule games, only understandable through the game” (Rakoczy and Tomasello 2007, 19).

We can try to find a third position between our two:

If we follow Paul Harris (2000) we have two hypothesis about the relation between reality and pretence²:

- a. *In the case of pretend play, children do not set up a contrast between an imaginary event and an actual event*
- b. *Pretend events are not set up as departures from actual events*

We have a good reason to agree with him about the first hypothesis and a plausible hypothesis to reject the second.

Contrast first. Why there is not real contrast? A generic notion of symbol doesn't help, but I think we can recognize that in our examples of pretence there is something that can be called iconic. According to Peirce a sign is an icon when it “may represent its objects mainly by its similarity” (Peirce 1931-1958, 2.276).

The iconic device may possess certain elementary iconic markers and sometimes a minimal resemblance is due to the fact that the iconic sign, even though different in shape from its object, performs the same function. We can think just a stick which qualifies as a horse because one can proudly ride it. The only aspect that the stick has in common with a real horse is that it can be straddled. Hence the child has rendered emergent one of the functions permitted by the horse. Now, contrary to what is sometimes said, communication need not come into this process at all³. He may not have wanted to show his horse to anyone. It just served as a focus for his fantasies as he galloped along.

But our stick can become the icon of a horse, a sceptre, or a sword. In all these objects recurs the same element, that is the feature of linearity (vertical or horizontal). But we haven't a case of imitation. Insofar as the vertical quality is a feature of both a stick and a sword it is the same verticality. Then, a linear dimension has been used as an expressive feature in order to substitute for the linear dimension that equally characterizes a horse as such. Or, in other words, a part of the referent is used as a sign-vehicle. We can quote Kant. In the first *Kritik* space, like time, is a pure intuition, the elementary form that we confer upon experiential data so as to be able to perceive them and place them within the categories. Therefore, verticality and horizontality are the intuitive mode within which we frame our perceptions, not intellectual abstractions. The spatial dimensions are not an intellectual construction, but the structural conditions for a possible object, and as conditions they

² In her review, Tamar Gendler defines Harris' work “an extraordinarily comprehensive and informative book surveying a tremendous range of empirical psychological work on imagination in children” (2002, 414). Gopnik (2009, ch. 1) is even more generous with Harris' work.

³ Here I agree with Ernest Gombrich (1951). It seems that Searle's insistence on the inescapable role of language goes beyond what is justified by the facts. Early *pretence* is just an example of this problem.

may be reproduced, equal to themselves, in varied circumstances. Whereas geometrical figures are already objects built under the framework of such conditions, and they cannot be reproduced as equal to themselves, but only as abstractions similar to previous constructions. This doesn't stop the stick from standing for the horse, but this happens because the linearity of the stick is not a construct but a condition of every other possible construction, and thus an intuitive artifice able to determine a space. We can concede that a toy horse or a stick is a sign, but we need to better distinguish the *imitans* from the *imitatum*, that which stands for something from that for which something else stands. So, according to this analysis, iconism makes no appearance, and, as they say in semiotics, we have only intrinsically coded acts. So, to try to better understand the nature of *pretence* we are allowed to put aside symbolism.

We can consider the second argument: a possible hypothesis about possible departures from reality.

Episodic memory involves re-experiencing situations. Kant argued that for the experience of objects (inner or outer) to be possible at all space and time must precede and structure all experiences of outer and inner states. Our experience must be grounded in space and time as pure "forms on intuition". If experience has a spatial content and if episodic memory is re-experiencing then episodic traces must have spatial content. The claim is that for there to be episodic memory in any creature, what makes it an *experience* is something non-conceptual. According to Robert Hanna "The thesis of *Non-Conceptualism* about mental content says that representational content is neither solely nor wholly determined by our conceptual capacities, and that at least some contents are both wholly and solely determined by non-conceptual capacities and can be shared by human and non-human animals alike" (Hanna 2008, 42). Kant developed in the *Transcendental Aesthetic* the idea that space and time are a priori subjective forms of sensibility. However, sensibility was supposed to include not only sense perception but also the "inner sense" imagination, pleasure, pain, desire. Infact, we should think of the representation of space and the representation of time as the necessary a priori subjective forms of egocentrically centered human and non-human animal embodiment. Following these premises we claim that memorial re-experiencing is a form of inner sense, akin to the imagination. In the case of *pretence* I think that the non-conceptual image might be identified with what we have seen as a minimum image that is still sufficient to release a specific reaction.

Imagining intends absent objects; perceiving intends present objects. Same objects, different intentional relation. So, is it possible to see something and, in the same time, to image something else that negates it? Here with "to image" I simply mean an episode of imagery, a mental image. Or, starting from a more *basic* question, is it possible to see and to image the same thing? For example, while we are looking at our mother we can try to visualize her face, and we need the same content in the very same way. But this exercise is not easy at all. It is known that there is overlap in the regions of the brain that are activated in seeing and visualizing. According to Kosslyn the same cerebral mechanism in our neuroanatomy must be involved, the Visual Buffer. In *Zettel*, Wittgenstein says "while I am looking at an object I cannot imagine it" (§621). This means that I cannot imagine the very object I am looking at. I can surely be looking at my mother from the back, not even realize I am looking at my mother, and still imagine her from the front. *De re* seeing does not prevent *de dicto* imagining. So our minimum image is projected into the real situation and it helps to imagine alternative metaphysically possible situations, that is real departures. I suggest that the child who pretends productively is tracking the world. He is responding to the world in something like the way one responds to the world with one's beliefs. Rules of pretend-inference have a robustly objective status, as demonstrated by Greg Currie (1998). On this view *pretence* stands to pretending as truth stands to beliefs.

REFERENCES

- Abelev, M. and Markman E. (2006), “Young Children’s Understanding of Multiple Object Identity: Appearance, Pretense and Function”, *Developmental Science*, 9(6), 2006, pp. 590-596;
- Bates, E., Benigni, L., Bretherton, I., Camaioni, L., and Volterra, V. (1979), “Cognition and Communication from Nine to Thirteen Months: Correlational Findings”, in E. Bates (Ed.), *The Emergence of Symbols: Cognition and Communication in Infancy*, Academic Press, New York, 1979;
- Billman, D. (1992), “Modeling Category Learning and Use”, in Burns B. (ed.), *Percepts, Concepts, and Categories. The Representation and Processing of Information*, North-Holland, Amsterdam, 1002;
- Currie, G. (1998), “Pretence, Pretending and Metarepresenting”, *Mind and Language*, 13, 1998, pp. 35-55;
- Gelman, S.A. and Byrnes, J.P. (eds.) (1991), *Perspectives on Language and Thought*, CUP, Cambridge, 1991;
- Gendler Szabó, T. (2002), “Review”, *Mind*, Vol. 111, No. 442, 2002, pp. 414-418;
- Gombrich, E. (1951), “Meditations on a Hobby Horse”, in *Meditations on a Hobby Horse and Other Essays on the Theory of Art*, (1963), Phaidon, London, 1951;
- Gopnik, A. (2009), *The Philosophical Baby*, Farrar, Straus and Giroux, 2009;
- Hanna, R. (2008), “Kantian Non-Conceptualism”, *Philosophical Studies*, 137, 2008, pp. 41-64;
- Harris, P. (2000), *The Work of the Imagination*, Wiley-Blackwell, Oxford, 2000;
- Kosslyn, S.M. (1994), *Image and Brain: The Resolution of the Imagery Debate*. MIT Press, Cambridge, MA, 1994;
- Leslie, A.M. (1987), “Pretense and Representation: The Origins of 'Theory of Mind'”, *Psychological Review*, 94, 1987, pp. 412-426;
- Lillard, A. (2001), “Pretend Play as Twin Earth: A Social-Cognitive Analysis”, *Developmental Review*, 21, 2001, pp. 495-531;
- Peirce, C.S. (1931-1958), *Collected Papers*, Harvard University Press, Cambridge;
- Rackoczy, H. (2008), “Taking Fiction Seriously: Young Children Understand the Normative Structure of Joint Pretence Games”, *Developmental Psychology*, 2008, p. 44;
- Rackoczy, H. and Tomasello M. (2007), “The Ontogeny of Social Ontology”, in S.L. Tsohatzidis (ed.), *Intentional Acts and Institutional Facts: Essays on John Searle’s Social Ontology*, Springer Verlag, Berlin, 2007;
- Scruton, R. (1974), *Art and Imagination*, Methuen, London, 1974;
- Searle, J.R. (1999), *Mind, Language, and Society. Philosophy in the Real World*, Weidenfeld and Nicholson, London, 1999;
- Searle, J.R. (2005), “What is an Institution”, *Journal of Institutional Economics*, 1(1), 2005, pp. 1-22;
- Searle, J.R. (2010), *Making the Social World*, OUP, Oxford, 2010;
- Wittgenstein, L. (1981), *Zettel*, (2nd. Ed.), G.E.M. Anscombe and G.H.V. Wright (Eds.), Blackwell, Oxford, 1981.

SESSION

3

SESSION 3

ETHICAL-POLITICAL LIFE

Massimo Reichlin (Università Vita-Salute San Raffaele, Milano)

The Neosentimentalist Argument Against Moral Rationalism:
Some Critical Observations

Christian Blum (Universität zu Köln)

Determining the Common Good:
A (Re-)Constructive Critique of the Proceduralist Paradigm

Roberta Sala (Università Vita-Salute San Raffaele, Milano)

Reasonable Values and the Value of Reasonableness.
Reflections on John Rawls' Political Liberalism

MASSIMO REICHLIN

Università Vita-Salute San Raffaele, Milano

reichlin.massimo@univr.it

THE NEOSENTIMENTALIST ARGUMENT AGAINST MORAL RATIONALISM: SOME CRITICAL OBSERVATIONS

abstract

On the basis of the empirical evidence concerning the role of emotions in moral judgments, new sentimental approaches to metaethics have been proposed. Nichols' theory of sentimental rules, in particular, associates the emphasis on emotive reactions to the relevance of a normative body of rules that guide our judgment on actions. According to Nichols, the emotive mechanism of concern explains the acquisition of the moral capacity and, together with the evidence on psychopaths and autistic children, shows the implausibility of a) moral rationalism, both as a conceptual and as an empirical thesis; b) motivational internalism; and c) moral objectivism. However, if we distinguish between i) the initial acquisition of morality in children and ii) the adult experience of it, we can see that to accept a central role of the emotive mechanisms in the first is not to have shown their centrality in the second. In particular, it is not possible to account for the normative theory in purely emotive terms, even though we accept that their emotive connotation favours the evolutive success of the norms. A moderate rationalist view, grounded on the notion of reflective endorsement and on the cooperation between emotions and rational capacities seems quite compatible with the empirical evidence and can justify plausible forms of internalism and objectivism.

keywords

Moral sentimentalism, moral rationalism, moral internalism, moral objectivism, Nichols.

1. Recent research in moral psychology and on the neural bases of moral judgment has highlighted the relevance of automatic reactions of an emotive kind in human morality. This research has empirical character, *i.e.* it aims to provide a reliable description of how our moral capacity *actually* works; according to some, it also has normative relevance, that is, it suggests how we should (or should not) use our moral capacities as well (Greene 2003 and 2008; Singer 2005). To say that neuroscientific research has such a normative bearing is however highly controversial, as it seems to imply a problematic derivation of “ought” from “is”: and some commentators have rightly insisted that in order to reach any such normative conclusions we must always presuppose certain moral intuitions, while the neuroscientific results are in fact doing no real work in the argument (Berker 2009).

It is much more plausible, however, to say that this research has some direct or indirect implication of a metaethical kind, that is, on our ideas concerning the nature of ethics. Metaethical propositions are not prescriptive, but descriptive: they aim to explain what *moral* judgments are, and whether they can be said to have any kind of truth-value. According to several authors, research in moral psychology and neuroethics justifies a radical revision of our traditional image of the moral capacity: specifically, a criticism of rationalist views, according to which our moral judgments are essentially the work of reason, and a vindication of a sentimentalist approach, that considers them the expression of our sentiments and affective reactions. Some explicitly say that there is empirical evidence for an emotivist conception of morality (Haidt 2001), according to which moral judgment belongs to moral sense, that is, to “an innate preparedness to feel flashes of approval or disapproval toward certain patterns of events involving other human beings” (Haidt and Joseph 2004, 56).

Although the data clearly do not show the truth of emotivism concerning the nature of *moral language*—that is, of the view holding that moral judgments have the *logical* function to express certain peculiar emotions (Joyce 2008)—they in fact seem to count as reasons in favor of a sentimentalist approach that, both from an explanatory and from an explicitly causal point of view, attributes a central role to emotive reactions and an altogether subsidiary role to rational reflection. In particular, Haidt’s social intuitionist conception aims to show that moral judgments are caused, in the great majority of cases, by the bursting in of certain emotions, while the arguments offered

in their favor are almost always *ex post* rationalizations, that is, clumsy attempts to give reasons to support our automatic responses: these attempts are generated in the context of a social request of justification and have no causal or explanatory role in generating the judgments in the first place.

This model seems to leave no space for altogether ordinary psychological processes, such as first-person solitary moral reflection and deliberation, and must consider all moral arguing as a lawyer's attempt to defend a moral truth that is simply accepted a priori. Moreover, it seems unable to account for the empirical fact that controlled "cognitive" processes are effective in (sometimes) contrasting the emotive responses, overwriting rational responses based on consequentialist computations (Greene *et al.* 2001, 2004 e 2008). In other words, by making reason a complete "slave of the passions", this emotivist view is in contrast with the empirical evidence that suggested the "dual-process model".

Shaun Nichols' model of the sentimental rules offers a much more complex formulation of the neosentimentalist approach: it not only establishes the disposition to feel certain emotive reactions, particularly with reference to others' sufferings, as the fundamental component of the moral capacity; it also stresses the importance of some normative perspective, that is, of a body of rules concerning the approval and disapproval of certain kinds of behavior (Nichols 2002a, 2004, 2008).

2.

Building on the data from several psychological experiments, Nichols argues for the following points: a) morality—and particularly the capacity to make *core moral judgments* (i.e., judgments that actions causing suffering to other people are wrong) and to distinguish between moral and conventional violations—is explained by the affective mechanism of *concern*; this accounts for the fact that three-year old children and autistic children are able to make core moral judgments while still lacking a developed theory of mind; b) the importance of this emotive mechanism, coupled with the empirical data concerning psychopaths and people suffering from focal lesions of the ventromedial section of the prefrontal cortex, shows the implausibility of moral rationalism; c) the fact that psychopaths often display an intact capacity for moral judgment shows that knowledge of the sentimental rules can be dissociated from the emotive mechanism: moreover, subjects suffering from focal lesions in the prefrontal cortex seem to be the living evidence of the theoretical possibility of amorality (Roskies 2003) and therefore show the fallacy of the conceptual internalism of the rationalistic view; d) the cultural fitness of emotively laden rules explains why certain rules do survive throughout human history and are largely accepted: this makes the objectivist explanation of moral "progress" offered by the rationalist view redundant. In other words, the partial and progressive overlap between moral codes can be explained by the similarity between our affective reactions; we simply happen to evolve such sentiments and there are no deeper moral facts that might vindicate moral objectivism.

In what follows, I will suggest some reasons to believe that the very fact that Nichols' sentimentalist view is much more plausible and attractive than Haidt's weakens his attack against moral rationalism, limiting its efficacy to radical or extreme forms of ethical rationalism.

In order to assess Nichols' basic points, let us distinguish between i) the initial experience of morality which is found in children and which emerged phylogenetically at a certain point in the evolution of the species *Homo* and ii) the much more complex experience of morality that is found in healthy adult humans. Nichols' research concentrates on level i): it aims to account for those basic judgments which he calls *core moral judgments*. A direct and easy way to defend moral rationalism would obviously be to deny that such judgments — that, according to many moral psychologists are found in three-year old children—are in fact authentic moral judgments. We might say that the moral faculty presupposes a much larger mastery of moral concepts and of complex moral reasoning and that, therefore, the data

3.

from moral psychology do not undermine moral rationalism. However, Nichols appropriately notes that children's moral judgments resemble the adults' ones in many ways; moreover, a considerable part of our daily normative experience is in fact based on those same mechanisms that are at work in core moral judgments.

It seems sensible, then, to accept these judgments as integral to the moral capacity and to find in them the essential elements contributing to moral judgments in adults. Nonetheless, the fact that core moral judgments depend on emotive reactions at level i) does not necessarily imply a sentimentalist explanation of level ii); that is, it does not demonstrate that emotive reactions to the others' suffering are the decisive element of the moral capacity, as it is possessed by adults, or that all moral judgments depend on the concern mechanism. In effect, although Nichols considers highly implausible that the concern mechanism should have a merely evolutionary role, so that adult moral judgment could function in its total absence, he deems just as much dubious the "on-line hypothesis", according to which the making of a moral judgment would always presuppose the activation of the affective mechanism: thanks to the normative theory, it seems very likely that adult moral judgments can often be made without any such activation.

Thus, even accepting that the emotive mechanism of concern is essential for the starting of the moral faculty (Nichols' point a), we need to have a clearer view of the relationship between this mechanism and the normative theory in order to see whether the model of the sentimental rules can count as a refutation of moral rationalism (Nichols' point b). Now, both elements seem necessary, according to Nichols: the ability to feel certain emotive reactions is a necessary presupposition of the moral judgment—since individuals with congenital frontal lesions never acquire the moral capacity; but the emotive reactions are not in themselves sufficient to characterize morality and to distinguish it from other normative spheres. In fact, on the one hand, in cases of natural disasters, or of human interventions causing suffering with a view to larger benefits (for example, in the medical field), we can observe emotive reactions similar to those linked to core moral judgments, but no moral judgment is generated. On the other hand, Nichols' experiments themselves show that prohibitions relative to disgusting behaviour, which are reinforced by strong emotive reactions, acquire the same weight of moral prohibitions, even though common sense clearly distinguishes between the two kinds of prohibitions. (Nichols 2002a). This suggests that the emotions are insufficient to distinguish the moral from the non moral sphere, and that the normative theory does not simply embody those emotive reactions. As a matter of fact, some non moral prohibitions relative to disgusting behaviour may be supported by negative emotions that are even stronger than those relative to some moral prohibitions, for example those tied to rules of justice and fairness in distribution, or those relative to promises: however, our normative theory considers the second as a clearly distinct kind of violation, and one much more serious than the first.

It seems therefore highly plausible to say that the normative theory that structures our emotive reactions reflects a relevant work of rational reflection as well. It is not the case that our set of rules simply expresses our emotions; it also reflects the work of rational reflection, a work which can partially shape our emotive reactions. This work manipulates specifically moral concepts—not only harm and physical suffering, but also injustice and unfairness, disrespect of others' dignity, humiliation of others, and the like—and on the basis of these concepts attributes a particular degree of importance to *specifically moral* violations; the acknowledgment of the peculiar importance of moral violations, on its turn, generates new emotions and sentiments that were not involved in core moral judgments. The very fact that Nichols' theory attaches a relevant role to the normative theory, therefore, shows that sentimental reactions are insufficient in order to give a complete explanation of morality and of its distinction from other forms of normativity. If this

is so, it is very likely that also in the initial acquisition of the moral capacity an essential role is played by the normative set of rules that enable to judge the appropriateness of feelings and actions; without the aid of the normative theory, children would never be able to distinguish between the moral from the non moral sphere, and would never learn the distinction between rules forbidding actions that are malicious or unfair and those forbidding actions that are simply disgusting.

A possible sentimentalist rejoinder is to insist that the normative theory is itself the product of our emotive reactions: it is the emotive mechanism that accounts for the rules we have, and the normative theory is simply the systematization our sentimental responses. Nichols himself proposes the “affective resonance hypothesis”, according to which the annexing of a rule into our normative system is considerably favoured by its cultural fitness, which in turn depends on our emotive reactions: that is, the more a rule is tied to a behaviour that in itself generates a remarkable emotive reaction, the more it is likely to be permanently adopted and handed down to future generations (2002b; 2004). However, we must note that the fact that being emotively reinforced favours the evolutionary success of a rule is quite compatible with the hypothesis that the normative theory is not a mere systematization of our spontaneous affective reactions. And Nichols himself does not deny that rules can be preserved from one generation to the next, even though they are *not* emotively laden. He writes: “*obviously* there are other important factors in cultural evolution. The hypothesis is only that affective resonance will be *one* of the factors that influence cultural evolution” (Nichols 2008, 270). It is clear, therefore, that the normative system cannot be explained exhaustively in terms of emotive causes: in particular, it is very plausible to suggest that some of our rules can be explained with reference to the aims that they serve, their favouring social cohesion, or their being supported by reasons that can be widely shared. In other words, lacking any proof of the fact that all our rules are based on an emotive sanction, it is likely that some (implicit or explicit) rational mechanism exists that, along with our emotive reactions, helps to account for our normative theory. It is this supplement of rational reflection that allows to distinguish between the moral rules and other rules, such as those relative to disgust, that are supported by emotive mechanisms as well.

4.

This shows that one of the factors explaining the adoption of the rules—even of those rules that are in fact supported by the emotive mechanism of concern—may well be the acknowledgement of objective reasons for their adoption, *e.g.* reasons consisting in the desire to protect people from unnecessary suffering and the like. In this picture, the emotive perception of others’ suffering would merely favour and support the perception of these reasons: the reasons themselves, however, could not be reduced to the emotions, being rather their causes. The fact that we have the normative code that we have would thus partly be explained by a mechanism of reflective endorsement, in which several normative reasons—including those that are brought about by our automatic emotive responses—are reviewed in a process of rational deliberation. If this is so, the core moral judgments themselves are generated by emotions that have been shaped by the normative theory governing the process of education; and this normative theory carries the traces of a complex cultural process. Moreover, it is clear that the moral experience of adult individuals decisively depends on a normative theory that is no longer simply received, but self-consciously accepted and critically discussed. At this level, moral judgments may be partially independent from the actual activation of emotive responses, as in fact they generally are: it is a datum of experience that many of our ordinary moral judgment are not associated to, nor in any way generated by, the on-line activation of moral emotions.

The conclusion to be drawn is that, while Nichols’ criticism seems to be effective against extreme rationalistic views, according to which reason provides a full explanation of morality, quite independently of emotive or sentimental reactions, it does not rule out more moderate rationalistic views. In fact, a) scientific evidence concerning the moral capacity of children and of psychopathic or

“acquired sociopathic” individuals shows that some basic emotive capacity is a prerequisite of “the moral point of view”, which is partly defined by the capacity to be emotionally tuned in with what happens to other people; and b) the normative theory that defines which emotions to have and which not, seems to be partly shaped by our emotive reactions. This, however, is compatible with a view according to which morality is, at root, the space of reasons, and in many cases (though not in all) moral judgments depend on some rational (and fallible) deliberation on the reasons that we have. More precisely, in young children’s core moral judgments the reference to reasons is implicit and indirect, since it is mediated by the normative theory, which, though represented internally, is not explicitly present to the consciousness of the judging individual. In adult individuals’ more complex judgments, on the other hand, there is almost always some direct reference to our reasons. In fact, according to Nichols himself, an individual who lacked any normative theory would not make real moral judgments, but would limit herself to express her emotions, as if uttering an interjection. In other words, contrary to XXth century emotivists, Nichols’ neosentimentalism does not treat emotions as in any way part of the content of moral judgments; emotions simply “play a role in leading us to treat as distinctive certain violations, including many of those we consider “moral”, like violations of harming others” (Nichols 2008, 263). This, however, is quite compatible with a rationalistic view according to which emotive reactions are primary sources of reasons to act, and moral judgments are the reflective endorsement of our best reasons, mediated by the normative theory.

5. Let us now move to Nichols’ thesis c), that is, the objection to motivational internalism: this is the view according to which there is a conceptual link between accepting a moral judgment and having some motivation to act accordingly. Nichols believes that empirical evidence on psychopaths shows the possibility to make moral judgments without acquiring the corresponding motivations in the least degree; however, it is far from obvious that the moral judgments of psychopaths should count as authentic. On this issue, Nichols is happy to rely on the empirical evidence of general opinion: he simply says that most people do believe that these moral judgments are authentic (Nichols 2002c e 2004). Even though we should accept this conclusion, it could hardly count as a decisive refutation of *rationalistic* internalism: according to this view, in fact, it is a conceptual truth concerning moral judgments that they provide reasons for action, which in turn provide motivations to act, *so far as we are rational* (Korsgaard 1986; Smith 2004 and 2008; Joyce 2008). And rationalistic internalism acknowledges the existence of many reasons that may restrict the motivational capacity of moral reasons: for one thing, weakness of the will, or the fact that, acknowledging the existence of a reason to do *x*, we acquire a motive to do it, but we are pulled by a stronger motive to do *y*; and other forms of practical irrationality, on account of which we fail to acknowledge the practical implications of what we sincerely declare to believe. Doubtless, psychopathy is itself a motive inhibiting the normal functioning of practical rationality; as long as they constitute blatant examples of irrationality, psychopaths cannot constitute an objection to the view that to accept a moral judgment and not to have any corresponding motivation is to be irrational, even if we should accept that their moral judgments are authentic.

On the other hand, to affirm the authenticity of these judgments is scarcely plausible. The very fact, stressed by empirical research, that practical reason is not a cold calculation of means and ends, but a cooperative enterprise between emotive responses and reflective processes, shows that the moral judgments made by psychopathic individuals are not authentic. As a matter of fact, they do not adequately distinguish between conventional and moral norms and therefore, in contrast to three-year old children, do not possess the moral capacity, according to Nichols’ criterion¹: we might say that they but quote socially widespread moral beliefs and rules. These patients lack the mechanism

1 Of course, one could deny that the capacity to distinguish between the moral and the conventional is constitutive of the moral faculty, as suggested by Roskies (2008).

of concern that helps understand the seriousness of moral violations and that motivates moral behaviour. In the standard terms of contemporary metaethical discussion, the moral judgments of psychopathic individuals are paradigmatic examples of an “inverted commas” use of moral language. In any case, the fact that psychopaths make real moral judgments might be an argument against “simple motivational internalism”, according to which moral judgment *guarantees* moral motivation, but not against the form of internalism peculiar to moral rationalism (Joyce 2008).

Similar considerations can be offered with reference to patients with ventromedial lesions of the prefrontal cortex. Since these patients do make moral judgments, but lack any motivation to act correspondingly, they have been considered concrete examples of amorality and direct proofs of the falsity of motivational internalism. As noted by Adina Roskies, they are “walking counterexamples to this internalist thesis” (Roskies 2003, 51). Again, how authentic these moral judgments can be, is a matter of discussion: and if they are, the fact that rational and emotive areas of their brain are unrelated justifies the conclusion that VM patients’ practical rationality is compromised; therefore, it is difficult to see how they could undermine the peculiar motivational internalism of the rationalistic view. On the other hand, it must also be stressed that there is a radical difference between making moral judgments in third and in first person: VM patients may be able to make “abstract” moral judgments from a spectator’s point of view, but are definitely unable to make first person moral judgments, since this would imply a commitment to act accordingly that they are unable to adopt. In this sense, although it may be incorrect to say that VM patients who are able to reason abstractly on moral dilemmas in Kohlberg’s style make “inverted commas” moral judgments—for they seem seriously convinced of the truth of what they say—it is nonetheless incorrect to say that they make authentic moral judgments. To do this, they should have the capacity to decide in first person, that is, to assume the agent’s perspective and to understand that action *x* is what is required *from me in these circumstances* (Kennett and Fine 2008).

One last word on the issue of moral objectivism (Nichols’ thesis *d*), or the idea that the sentimental rule approach and its story concerning the cultural fitness of emotively laden rules undermine moral objectivity. 6.

For one thing, it must be noted that moral rationalism is committed to moral objectivism, but not necessarily to moral realism: for example, a rationalist view such as Greene’s or Singer’s affirms the necessity to reformulate our evolutionary-biased normative theory in a consequentialist shape to reach moral objectivity, but does not embrace moral realism, nor justify any belief in objective moral facts (Singer 2005; Greene 2003; Greene 2008).

On the other hand, we cannot exclude that emotions might point out moral facts: they might not be contingent evolutionary facts, but means to acquire knowledge of some independently existing moral reality. Of course, the plausibility of this view depends on how such a “moral reality” is conceived: if to accept that there are moral facts is to say that, in each situation, there are objective moral reasons to act, to claim that the emotions we have are the “right” ones because they point out the moral facts is simply to claim that in the emotive reactions accompanying the consciousness of some other’s suffering or being treated unjustly, we have access to objective reasons to avoid that suffering or injustice: reasons, that is, that can be acknowledged and shared by everyone. Rational reflection can detect these reasons and decide whether to adopt them, that is, whether to transform them in maxims of action and possibly in general principles of our normative theory. This does not alter the fact that emotive reactions are major sources of such objective reasons that our reflection can sanction.

In other words, the rationalistic view affirming the practical character of reason does not necessarily (nor generally) maintain that moral statements refer to purely objective facts, utterly independent of any emotion or practical interest of the agent. It claims that they express the reasons for action that are highlighted both by our emotive reactions and by our rational reflection, while facing practical situations. Moral choice, therefore, expresses both our emotive and rational nature, namely, the embodied nature of our moral agency.

REFERENCES

- Berker, S. (2009), "The Normative Insignificance of Neuroscience", *Philosophy and Public Affairs*, 37, 2009, pp. 293-329;
- Greene, J. (2003), "From Neural 'Is' to Moral 'Ought': What Are the Moral Implications of Neuroscientific Moral Psychology?", *Nature Reviews Neuroscience*, 4, 2003, pp. 847-850;
- Greene, J. (2008), "The Secret Joke of Kant's Soul", in W. Sinnott-Armstrong (ed.), *Moral Psychology. Volume 3: Emotion, Brain Disorders, and Development*, MIT Press, Cambridge, Mass., pp. 35-79, 2008;
- Greene, J. et al. (2001), "An fMRI Investigation of Emotional Engagement in Moral Judgment", *Science*, 293, 2001, pp. 2105-2108;
- Greene, J. et al. (2004), "The Neural Basis of Cognitive Conflict and Control in Moral Judgment", *Neuron*, 44, 2004, pp. 389-400;
- Greene, J. et al. (2008), "Cognitive Load Selectively Interferes with Utilitarian Moral Judgment", *Cognition*, 7, 2008, pp. 1144-1154;
- Haidt, J. (2001), "The Emotional Dog and its Rational Tail: A Social Intuitionist Approach to Moral Judgment", *Psychological Review*, 108, 2001, pp. 814-834;
- Haidt, J. and Joseph, C. (2004), "Intuitive Ethics: How Innately Prepared Intuitions Generate Culturally Variable Virtues", *Daedalus*, 133, 2004, pp. 55-66;
- Joyce, R. (2008), "What Neuroscience Can (and Cannot) Contribute to Metaethics", in W. Sinnott-Armstrong (ed.), *Moral Psychology. Volume 3: Emotion, Brain Disorders, and Development*, MIT Press, Cambridge, Mass., 2008, pp. 371-394;
- Kennett, J. and Fine, C. (2008), "Internalism and the Evidence from Psychopaths and 'Acquired Sociopaths'", in W. Sinnott-Armstrong (ed.), *Moral Psychology. Volume 3: Emotion, Brain Disorders, and Development*, MIT Press, Cambridge, Mass., 2008, pp. 173-190;
- Korsgaard, C. (1996), "Skepticism About Practical Reason", in *Creating the Kingdom of Ends*, Cambridge University Press, Cambridge, 1996, pp. 311-334;
- Nichols, S. (2002a), "Norms with Feeling: Towards a Psychological Account of Moral Judgment", *Cognition*, 84, 2002, pp. 221-236;
- Nichols, S. (2002b), "On the Genealogy of Norms: A Case for the Role of Emotion in Cultural Evolution", *Philosophy of Science*, 69, 2002, pp. 234-255;
- Nichols, S. (2002c), "Is It Irrational to Be Immoral? How Psychopaths Threaten Moral Rationalism", *The Monist*, 85, 2002, pp. 285-304;
- Nichols, S. (2004), *Sentimental Rules. On the Natural Foundations of Moral Judgment*, Oxford University Press, Oxford, 2004;
- Nichols, S. (2008), "Sentimentalism Naturalized", in W. Sinnott-Armstrong (ed.), *Moral Psychology. Volume 2: The Cognitive Science of Morality: Intuition and Diversity*, MIT Press, Cambridge, Mass., 2008, pp. 255-274;
- Roskies, A. (2003), "Are Ethical Judgments Intrinsically Motivational? Lessons from 'Acquired Sociopathy'", *Philosophical Psychology*, 16, 2003, pp. 51-66;
- Roskies, A. (2008), "Internalism and the Evidence from Pathology", in W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, MIT Press, Cambridge, Mass., 2008, pp. 191-206;
- Singer, P. (2005), "Ethics and Intuitions", *The Journal of Ethics*, 9, 2005, pp. 331-352;
- Smith, M. (2004), *The Moral Problem*, Basil Blackwell, Oxford, 2004;
- Smith, M. (2008), "The Truth About Internalism", in W. Sinnott-Armstrong (ed.), *Moral Psychology, Volume 3: The Neuroscience of Morality: Emotion, Brain Disorders, and Development*, MIT Press, Cambridge, Mass., 2008, pp. 207-215.

CHRISTIAN BLUM

Universität zu Köln

chr-blum@gmx.de

DETERMINING THE COMMON GOOD: A (RE-)CONSTRUCTIVE CRITIQUE OF THE PROCEDURALIST PARADIGM

abstract

On the basis of a short summary of phenomenological aims and methods, this essay describes the present state of relationships between phenomenology and analytic philosophy, pointing out the progress done in the last years on the way of their rapprochement, after a long time of reciprocal scorn and misunderstandings. In the way of a presentation of the Phenomenology Lab and Center's present and future research program, it recalls some relevant chapters of past and present phenomenological research in Europe, and quite particularly in Italy. After discussing some aspects of contemporary debates in phenomenology and philosophy of mind, it attempts at establishing a convergent line of argument toward the assessment of an anti-reductive ontology of concreteness, or the life world.

keywords

Common good, democracy, proceduralism, metaethics

- The Renaissance of the Common Good**
1. For more than a decade, philosophers, political scientists and legal theorists have witnessed a “renaissance of the idea of the common good” (Offe 2012: 3) in political discourse: political agents such as parties, civic networks and courts increasingly refer to the concept of the common good in order to justify their actions or criticize the actions of political antagonists (Amitai 2004; Riordan 2008). The revitalization of this political concept is largely due to a growing discontent with the classical liberal model of political legitimacy advocated by theorists such as Rawls (1971, 2005) and Dworkin (1977). Contrary to classical liberalism, it is argued that legitimacy cannot be reduced to a just consideration of *individual* claim-rights on certain divisible goods (e.g. rights, societal positions, income, health services) but that it must as well account for the *collective* interests of society as a whole (Taylor 1995, 2003; Freeman 2000).

This renaissance of the common good, however, is also met with skepticism by many theorists: First, the very concept is considered inimical to a pluralistic democratic society. Assuming a normative entity such as the common good of a community, critics argue, amounts to prescribing a pre-existing “common way of life” (Kymlicka 1997, 226) to which all citizens must adhere. This common way of life would severely restrict the scope of admissible political decisions, factually rendering democratic decision-making, which allows all citizens – regardless of gender, religion, or sexual orientation – to assert their interests in an equal fashion illegitimate. Second, it is questionable whether the concept really “points to a clearly definable range of considerations in support of a policy” (Barry 1964, 1). Rather, the suspicion suggests itself, that appeals to the common good are, as Barry aptly puts it, nothing but “a handy smoke-screen” (Barry 1964, 1) employed by political agents to ensure support for their particularistic goals; apart from this merely rhetorical function the concept is too vague to serve any normative role in political discourse.

If political philosophy is to provide a sensible definition of the common good – thus taking account of the discontent with classical liberalism by expanding its model of political legitimacy –, it must deal with this criticism by showing, first, that the concept of the common good, properly understood, is not inimical to pluralistic democratic society; and, second, that it serves to identify a specific range of policies and hence fulfills an important role in political discourse. In the following, I will argue

that the paradigm of *proceduralism* which is advocated by several contemporary democratic theorists comes close to fulfilling these two requirements but must be modified in significant respects to provide a convincing definition of the common good.

The central thesis of proceduralism which, in its modern form, harks back to the works of Schubert (1960) and Fraenkel (1964, 1991) is the following:

The common good consists in the output of a political system whose procedures meet specific normative and functional standards of adequacy.

Normative standards of adequacy are generally defined in terms of democratic conditions that guarantee the fair and equal participation of all citizens in collective decision-making processes (Offe 2012, 16); the functional adequacy of a political system is determined by its responsiveness to the citizens' asserted interests and by its effectiveness and efficiency in implementing them through policies (Putnam 1992, 63-73)¹. The basic model of proceduralism is best understood as an input-output structure: the input into the political system consists in the citizens' interests that are asserted via participatory procedures (elections, referendums, petitions etc.). The asserted interests are received by the system's institutions and implemented through policies (laws, budget resolutions, executive decisions etc.) which – taken as a whole – constitute the system's output.

Two aspects of the proceduralist paradigm are of special significance for its assessment: First, proceduralists claim that a sensible theory does not require any substantive concretization of the common good itself (e.g. in the form of a list of collective goods) but must only specify *formal criteria* for political procedures; if the latter are fulfilled by a political system, that system's output necessarily constitutes the common good. This thesis is encapsulated in Gutman's & Thompson's claim that "once the right procedures are in place, whatever emerges from them is right" (Gutman and Thompson 2004, 24). Consequently, proceduralists deny that the output of a system can be criticized as detrimental to the common good by any criteria other than the standards of normative and functional adequacy themselves.

Second, proceduralists tend to *argue only negatively* for the plausibility of their paradigm by claiming that any substantive, i.e. non-procedural, concept of the common good is unsustainable: If political theory sought to provide a substantive definition of the common good (such as "the common good consists of collective goods x, y, z and must be realized by actions p, q, r"), it would undermine the legitimacy of democratic governance by pre-empting the output of collective decision-making (Fraenkel 1964, 1991, 300f.; Gutman and Thompson 2004, 25). Put in less technical terms: If there was already a substantive definition of the common good (offered by political theorists), there would be no need for citizens to engage in voting or public deliberation; the reasonable consequence would instead be a paternalistic governance by political theorists, because the latter would ensure that the common good is indeed realized. Since this inference, however, is clearly unacceptable for ethical reasons, substantive theories must be rejected – and proceduralism remains as the sole alternative.

This brief sketch illustrates the advantages as well as the indeterminacy of the proceduralist paradigm. On the one hand, proceduralism appears to be the ideal candidate for fulfilling the requirements of a sensible theory of the common good: Not only does it provide a definition that

¹ The question of how to specify exactly these criteria and how to balance them against each other in cases of conflict is of course a pressing matter for normative democratic theory that has engendered a vital debate. Since I am interested in assessing the plausibility of the proceduralist paradigm as such – and not in its concretizations – I shall not delve into this issue; for an overview see Christiano (2006).

is compatible with a pluralistic democratic society by positing that the common good must be understood as the output of fair and functional democratic procedures in which each citizen – irrespective of gender, sexual orientation, or religion – has an equal say, it also allows the identification of a specific range of policies, since for determining whether a political action is indeed conducive to the common good, we must simply ascertain whether it has been authorized by fair and functional democratic procedures. On the other hand, the fact that proceduralists argue only negatively for the plausibility of their paradigm gives reason for concern about a *positive corroboration* of proceduralism. Even if the negative argument was sound and substantive theories of the common good were unsustainable, there would remain the possibility that the proceduralist paradigm itself contains (tacit) premises that are indefensible. In the following, I will hence reconstruct the normative premises to which proceduralism ought to be rationally committed and determine, whether those serve as a convincing basis for the central thesis of proceduralism.

3.
The Subjectivist
Foundation of
Proceduralism:
Exercising
Popular
Sovereignty

The most promising normative basis from which to reconstruct the proceduralist paradigm is, in my view, the theory of *metaethical subjectivism* (Wiggins 1987, 1998; Sumner 1995)². This theory is based on the thesis that the ethically good (in our case, the common good) consists in the realization of the actual interests of individuals (in our case, the citizens of a political community). The notion of *interest* refers to a class of conscious mental states in which persons desire, prefer or seek certain states of affairs that are considered valuable by them. Since interests, thus understood, are accessible introspectively by their bearers, they can be intersubjectively communicated in the form of truth-apt propositional statements and asserted via participatory procedures. The punchline of metaethical subjectivism is that states of affairs derive their value exclusively from the fact *that* they are valued by individuals. This implies, in turn, that any statement according to which a collectively valued state of affairs is unethical is not simply false but meaningless.

The appeal of metaethical subjectivism with regard to the personal level is that it takes persons seriously as autonomous authors of their individual welfare and closes the door on paternalistic interventions (Arneson 2005). With regard to the political level, metaethical subjectivism helps to make sense of the notion of popular sovereignty that is central to democratic societies. According to this notion, democratic citizens must be understood as agents who autonomously shape the welfare of their community on the basis of their collective will³. The subjectivist foundation of proceduralism can be summarized by a principle that I call the sovereignty principle:

sovereignty principle: the citizens of a political community possess defining power over the common good.

By the expression “defining power” I simply mean that the common good of a given society is defined by the interests of that society’s citizens and that the latter are conceptually and metaphysically

² This does not mean, of course, that proceduralism necessarily implies metaethical subjectivism. Following Condorcet (1785, 1972), one might also argue for an objectivist version of proceduralism. This approach would have to assume the following premises: (1) the common good is constituted by objective ethical criteria that are epistemically accessible; (2) political systems serve the purpose of identifying said criteria and implementing them through policies; (3) it is possible to specify formal standards of adequacy which guarantee the reliable identification and implementation of said criteria. From these premises, it follows that the common good necessarily consists in the output of an adequate system because that system’s procedures are sufficiently reliable so as to always identify and implement the criteria that constitute the common good. I think, however, that this account is vastly implausible. For one, I am skeptical of any theory which holds that the notion of welfare (be it individual or collective) can be defined in purely objective terms, but I shall not dwell on this issue here. It should be evident that the third premise is untenable. Apart from logical inference-forms there is no epistemic procedure that guarantees success. But this is exactly the assumption which defenders of the objectivist approach must hold, since otherwise the above-mentioned conclusion would not follow.

³ This conception of popular sovereignty that seems so familiar and almost trivial to us nowadays harks back to the writings of Rousseau, particularly to his *Social Contract* (Rousseau 1762, 2008).

prior to the common good. Accordingly, the positive value of a certain political action for the furthering of the common good is a function of the respective valuing of that action by the members of society themselves.

Even though the sovereignty principle constitutes the foundation of the proceduralist paradigm as I understand it, it is – taken by itself – underdetermined. The reason is that one cannot assume a priori that all citizens always and immediately agree on the states of affairs that are conducive to the common good. On the contrary, it is more realistic to assume a high degree of dissent on this issue in modern pluralistic societies. The sovereignty principle thus only stipulates *that* the citizens possess defining power over the common good; it does not specify, however, *how* this defining power is to be realized and how to deal with the problem of dissent. This question is answered by a second principle which I call the procedural principle:

procedural principle: the citizens of a political community exercise their defining power over the common good through political procedures that allow them to assert their interests and that meet certain normative standards of participatory equality and functional standards of effectiveness and efficiency.

The introduction of procedures that facilitate the citizen's defining power over the common good gives rise to two new questions that are reflected in the above mentioned standards: first, there is the *normative* question of how much weight to accord to the respective interests of the various citizens that constitute the community. Second, there is the *practical* question of how to design the procedures to ensure that the citizens' asserted interests are indeed implemented through policies.

The normative standards of participatory equality can be justified in two steps. First, we can posit as a default assumption that the interests of each citizen must be considered equally in the political process of determining the common good. Christiano argues with great force for the inevitability of this assumption by stating that “[t]his equality proceeds from the importance of interests as well as the separateness of persons. [...] No one's interests matter more than anyone else's.” (Christiano 2004, 269) This premise, however, does not necessitate a claim on participatory equality by all citizens. If one assumes, realistically, that some citizens possess more political knowledge than others, one might well argue that the equal consideration of interests is best served by a regime of paternalistic experts (Estlund 2008, 30f.). To arrive at the standard of participatory equality it is thus necessary to provide, in the second step, a sound argument against paternalism. The most convincing argument proceeds from the anthropological fact that most persons are struck with a “cognitive bias” (Christiano 2001, 205) in regard to their own interests. Even citizens who are benevolently inclined towards others and who possess superior knowledge are at risk of misrepresenting the desires of others because of their necessarily limited individual perspective. To obviate these negative effects of cognitive bias on the determination of the common good it must be ensured that all citizens receive the same rights to assert their interests through political participation.

Compared to the normative standards of equality, the justification of the standards of functionality requires far less argumentative effort. If political procedures are to serve the purpose of facilitating the exercise of defining power, they must ensure that the citizens' interests do not simply “seep away” or are distorted by political decision-makers but are indeed realized through adequate policies. They must be, in other words, effective. Furthermore, one must assume that political procedures generally operate under conditions of relatively scarce material and temporal resources. This leads to the requirement that procedures must make optimal use of the available resources in the implementation of policies, i.e., they must be efficient.

Now that the sovereignty principle and the procedural principle have been established, there is a normative foundation to substantiate the central thesis of the proceduralist paradigm: the common good consists in the output of a normatively and functionally adequate political system *because* it represents the fair, effective and efficient realization of the citizens' asserted interests which, in turn, determine the content of common good itself. On this reading, the political system fulfills the function of realizing the citizens' defining power over the common good, and accordingly its output represents the result of the exercise of that defining power.

- 4. A Critique of the Proceduralist Paradigm: the Problem of Error**
- Although the central thesis of proceduralism can be positively corroborated by the theory of metaethical subjectivism it is, as I now shall argue, open to severe criticism – a fact which necessitates a critical re-evaluation of the proceduralist tenet that the common good can be determined exclusively on the grounds of formal procedural standards. The flaw of the proceduralist paradigm based on metaethical subjectivism is that it cannot explain how citizens could be in *error* about the common good.

The decisive argument has already been put forward by Rousseau in the *Social Contract* where he states: “Our will is always for our own good, but we do not always see what that is.” (Rousseau 1762, 2008, 66). Rousseau’s point is that citizens’ political interests may be fundamentally misguided and hence, when realized, not conducive to the common good but irrelevant or even detrimental to it. The reasons for such misguided interests are manifold: citizens may be misinformed about politically relevant states of affairs, they may lack significant information, or they may draw inconsistent inferences from accurate information (Schumpeter 1976, 262; Caplan 2007). Thus, it appears intuitively plausible that even political actions which are collectively authorized by fair and functional democratic procedures may fail to ascertain the common good because their underlying interests are deficient in the above mentioned respects. In the terminology of the input-output model from section 2, we might say that in such cases a deficient input generates a deficient output.

The flaw of proceduralism, as I have reconstructed it, is that it categorically denies this possibility. It claims that in order to determine the common good we must simply implement and observe the standards of fair equality and functionality because the latter guarantee an adequate realization of the citizens’ interests which, in turn, are deemed constitutive of the common good. If it is, however, implausible to assume (for the above reasons) that the citizens’ interests *necessarily* constitute their common good, then the implementation and observance of the procedural standards guarantees by no means that their output is indeed conducive to the common good.

Now the defenders of proceduralism might resort to what I call an *idealization strategy*. They might say: “Of course we do not mean that that common good consists in the output of a system whose members are ill-informed about relevant political matters. That would be absurd! We rather mean that it consists in a system’s output as it would result if its members were fully informed and rational.” In my view, this strategy which has been employed e.g. by Dahl (1989) and Cohen (1989) constitutes a pyrrhic victory for proceduralism. If one substitutes the *actual* citizens’ defining power over the common good by the defining power of *hypothetical* agents (who possess all the skills and information that their actual counterparts lack), one severs the proceduralist paradigm from the normative foundation that has lent it its plausibility in the first place – the notion of popular sovereignty. By arguing that the common good ought to be understood as the output of procedures that are observed by hypothetical agents, the defenders of proceduralism deny that the actual citizens are fit to determine the common good themselves via democratic decision-making. From this point it seems to be only a short way to justifying a paternalistic regime, which is a conclusion that proceduralists have firmly sought to avoid in the first place.

It appears that the defenders of proceduralism are caught in a dilemma: either they stick with their central thesis, but then they must deny, rather implausibly, that actual citizens could be in error about the common good. Or they supplant the actual citizens' defining power over the common good with the theoretical construct of the input of counterfactually fully informed and rational agents, but then they are on a straight path towards paternalism and forego the initial advantages of their approach. I find both alternatives equally unsatisfactory which is why I will propose a significant modification to the proceduralist paradigm in the final section.

Since traditional proceduralism as based on the theory of metaethical subjectivism fails and since I do not see another way to corroborate its central thesis, I propose a modification to the paradigm that restricts the original tenet described in section 2 according to which the common good can be defined exclusively on grounds of formal, procedural standards. The alternative definition that I suggest is the following:

The common good consists in the output of a political system if and only if (a) that system's procedures meet specific standards of normative and functional adequacy and (b) that output does not violate specific substantive and objective standards of adequacy that are determined by experts.

The difference to traditional proceduralism is that in my qualified version the procedural authorization of a political action is only a necessary but not a sufficient condition for its conduciveness to the common good. This modification allows the preservation of the citizen's defining power over the common good (though in a bounded sense) while solving the problem of error that I have addressed in section 4.

Let me start by saying what I consider to be the valuable point of traditional proceduralism as I have reconstructed it: proceduralists posit correctly, as I think, that the populace of a democratic community must be acknowledged as a group of agents who autonomously shape their collective welfare via democratic decision-making. In my view, the flaw of proceduralism is not that it acknowledges the citizens' defining power over the common good; the problem is that it *overstates* the point by making the common good exclusively dependent on the citizens' asserted interests and thus denies the possibility of error. Consequently, the challenge consists in developing a theory that takes into account the citizens' defining power *and* acknowledges that they might still be in error about the common good. This twofold requirement, however, cannot be met by the defenders of traditional proceduralism as they find themselves caught in the dilemma that I have sketched at the end of section 4.

In my view, the only rational solution to this problem is to scale down the status of democratically exercised defining power from a necessary and sufficient condition to a merely necessary condition: on this account, it is necessary for an action's conduciveness to the common good that it reflects the citizens' subjective interests which are asserted via normatively and functionally adequate democratic procedures; it is, however, not sufficient, since those interests may still be ill-conceived in regard to certain standards that transcend the citizens' subjective perspectives. Such standards must fulfill two conditions. First, they must be *objective* in the sense that they are independent of the citizens' actual interests as well as epistemically accessible, so that they can serve to identify and criticize interests that are misguided (i.e. irrelevant or detrimental) with regard to the common good. Second, they must be *substantive* (i.e. non-procedural), for the simple reason that they must serve the purpose of criticizing the value of outputs that otherwise satisfy all standards of procedural adequacy.

Now the question emerges as to how to concretize these objective and substantive standards of

5. Qualified Proceduralism: Introducing Substantive Standards

adequacy beyond the rather abstract indications that I have given here. I think, however, that this is not a task for political philosophy but for empirical experts who deal with specific areas of political interest. Here is a short example to support my point. In recent years, there has been a heated debate in Germany about the legal instrument of preventive detention (Flaig 2009). The supporters of a stricter criminal law who held the majority in both legislative chambers argued that it would serve the common good to pass a set of laws that allow the indefinite detention of imprisoned criminals if psychological evaluation indicates that they will pose a threat to the public once they are released. According to traditional proceduralism, the respective laws, which were eventually passed in 2004 with great public approval, would have had to be considered conducive to the common good for the simple reason that they were authorized by fair and functional democratic procedures.

Since then, however, empirical inquiry by criminologists has shown that the proponents of said laws have based their argument on a misconception: in fact, the psychological evaluation of imprisoned persons does not provide an accurate prediction of their future criminal conduct (Kinzig 2008, Alex 2010). Thus, it must be assumed that the vast majority of persons who have been indefinitely detained on the basis of these laws do not pose a threat to the public after all and that the respective legal instrument has no significant impact on collective welfare. In the terminology of my version of qualified proceduralism this insight constitutes the identification of an objective and substantive standard by experts, since it serves to identify and criticize the output of fair and functional procedures as irrelevant to the common good.

Even if this example supports my argument there is one last problem that must be addressed. In section 2 I sketched the negative argument employed by proceduralists in order to buttress their account. This argument states that traditional proceduralism is the only ethically acceptable theory of the common good, since any account that provides substantive criteria of the common good pre-empts the outcome of democratic decision-making and leads to paternalism. I believe, though, that this argument does not undermine my account of qualified proceduralism. It would only do so if the stipulated objective and substantive standards were considered to be necessary and sufficient for the determination of the common good. However, I have argued that in order for an action to be conducive to the common good it must – apart from not violating substantive and objective standards – also be authorized by the populace itself, otherwise it would not reflect the general will of the sovereign citizens. Consequently, democratic decision-making fulfills an indispensable function in my account of the common good; I merely hold that it may be based on misguided interests and hence requires examination by experts who possess superior knowledge in the respective areas of political concern.

REFERENCES

- Alex, M. (2010), *Nachträgliche Sicherungsverwahrung: ein rechtsstaatliches und kriminalpolitisches Debakel*, Felix-Verlag, Holzkirchen, 2010;
- Amitai, E. (2004), *The Common Good*, Polity Press, Cambridge, M.A., 2004;
- Arneson, R.J. (2005), "Joel Feinberg and the Justification of Hard Paternalism", *Legal Theory* 11 (3), 2005, pp. 269-284;
- Barry, B. (1964), "The Public Interest", *Proceedings of the Aristotelian Society*, Supplementary Volumes 38, 1964, pp. 1-18;
- Caplan, B. (2007), *The Myth of the Rational Voter: Why Democracies Choose Bad Policies*, Princeton University Press, Princeton, 2007;
- Christiano, T. (2001), "Knowledge and Power in the Justification of Democracy", *Australasian Journal of Democracy* 79 (2), 2001, pp. 197-215;
- Christiano, T. (2004), "The Authority of Democracy", *The Journal of Political Philosophy* 12 (3), 2004, pp. 266-290;
- Christiano, T. (2006), "Democracy", in Zalta (Ed.), *Stanford Encyclopedia of Philosophy*, The Metaphysics Research Lab, Stanford, <http://plato.stanford.edu/entries/democracy>;
- Cohen, J. (1989), "Deliberation and Democratic Legitimacy", in Hamlin and Pettit (Eds.), *The Good Polity: Normative Analysis of the State*, Basil Blackwell, Oxford, 1989, pp. 17-34;
- Condorcet, J.A. (1785, 1972), *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*, Chelsea Publ. Co, New York;
- Dahl, R.A. (1989), *Democracy and its Critics*, Yale University Press, New Haven-London, 1989;
- Dworkin, R. (1977), *Taking Rights Seriously*, Harvard University Press, Cambridge, M.A., 1977;
- Estlund, D.M. (2008), *Democratic Authority: a Philosophical Framework*, Princeton University Press, Princeton, 2008;
- Flaig, A. (2009), *Die nachträgliche Sicherungsverwahrung*, Peter Lang Verlag, Frankfurt a.M, 2009;
- Fraenkel, E. (1964, 1991), *Deutschland und die westlichen Demokratien*, Suhrkamp, Frankfurt a.M;
- Freeman, S. (2000), "Deliberative Democracy: a Sympathetic Comment", *Philosophy and Public Affairs* 29 (4), 2000, pp. 371-418;
- Gutman, A. and Thompson, D. (2004), *Why Deliberative Democracy?*, Princeton University Press, Princeton-Oxford, 2004;
- Kinzig, J. (2008), *Die Legalbewährung gefährlicher Rückfalltäter: zugleich ein Beitrag zur Entwicklung des Rechts der Sicherungsverwahrung*, Duncker and Humblot, Berlin, 2008;
- Kymlicka, W. (1997), *Contemporary Political Philosophy: an Introduction*, Oxford University Press, Oxford-New York, 1997;
- Offe, C. (2012), "Whose Good is the Common Good?", *Philosophy and Social Criticism*, 2012, pp. 1-20;
- Putnam, R.D. (1992), *Making Democracy Work: Civic Traditions in Modern Italy*, Princeton University Press, Princeton-Oxford, 1992;
- Rawls, J. (1971, 2005), *A Theory of Justice*, Harvard University Press, Cambridge, M.A.;
- Riordan, P. (2008), *A Grammar of the Common Good: Speaking of Globalization*, Continuum, London-New York, 2008;
- Rousseau, J.J. (1762, 2008), *The Social Contract*, transl. by C. Betts, Oxford University Press, Oxford-New York;
- Schubert, G. (1960), *The Public Interest: a Critique of the Theory of a Political Concept*, Free Press of Glencoe, Glencoe, 1960;
- Schumpeter, J.A. (1942, 1976), *Capitalism, Socialism and Democracy*, Harper & Row, New York;
- Sumner, L.W. (1995), "The Subjectivity of Welfare", *Ethics* 105 (4), 1995, pp. 764-790;
- Taylor, C. (1995, 2003), "Cross-Purposes: the Liberal-Communitarian Debate", in Matravers and Pike (Eds.), *Debates in Contemporary Political Philosophy: an Anthology*, Routledge, London-New York;
- Wiggins, D. (1987, 1998), *Needs, Values, Truth: Essays in the Philosophy of Value*, Oxford University Press USA, New York.

ROBERTA SALA

Università Vita-Salute San Raffaele, Milano

sala.roberta@univr.it

REASONABLE VALUES AND THE VALUE OF REASONABLENESS. REFLECTIONS ON JOHN RAWLS' POLITICAL LIBERALISM

abstract

This paper aims to question the idea of reasonableness in Rawls' account of political liberalism. My point is that reasonableness as the moral basis of the liberal society provides restrictions for differences – be they philosophical, moral, religious, cultural – to be included in the liberal society. Notwithstanding Rawls' attempt to expand political boundaries and to include those people who do not share moral liberal justification to justice as fairness, reasonableness selects “values holders” and assigns to the so-called “reasonable” a place in the political debate. The others, the “unreasonable”, are expected to become reasonable; alternatively, they would be paid control or even coercion in all the circumstances in which they should represent a risk for political stability. I believe that Rawls gives an incomplete account of unreasonableness: there may well be persons who are not “reasonable” in Rawlsian terms but who do not necessarily represent a danger for the just society. By the fact that they do not endorse values as freedom and equality in which fair cooperation is grounded, we cannot infer that they will necessarily try to violate the terms around which cooperation is structured by imposing their values on others. I proceed as follows: a) I detail the Rawlsian political turn in defending justice as fairness; b) I focus on the idea of reasonableness as the core of political liberalism; c) I defend the thesis that political liberalism needs to revise the idea of unreasonableness if it aims to deal with pluralism as a social and political fact.

keywords

Legitimacy, toleration, political justification, agreement, pluralism

This paper aims to question the idea of reasonableness in Rawls' account of political liberalism. My point is that reasonableness as the moral basis of the liberal society provides restrictions for differences – be they philosophical, moral, religious, cultural – to be included in the liberal society. Notwithstanding Rawls' attempt to expand political boundaries and to include those people who do not share moral liberal justification to *justice as fairness*, reasonableness selects “values holders” and assigns to the so-called “reasonable” a place in the political debate. The others, the so called “unreasonable”, are expected to become reasonable; alternatively, they would be paid control or even coercion in all the circumstances in which they should represent a risk for political stability. I believe that Rawls gives an incomplete account of unreasonableness: there may well be persons who are not “reasonable” in Rawlsian terms but who do not necessarily represent a danger for the just society. By the fact that they do not endorse values as freedom and equality in which fair cooperation is grounded, we cannot infer that they will necessarily try to violate the terms around which cooperation is structured by imposing their values on others. I proceed as follows: a) I detail the Rawlsian political turn in defending *justice as fairness*; b) I focus on the idea of reasonableness as the core civic value of political liberalism; c) I defend the thesis that political liberalism needs to revise the idea of unreasonableness if it aims to deal with pluralism as a social and political fact.

1. **The Political Turn of Justice as Fairness** I start by recapitulating the “political turn” from *A Theory of Justice* (Rawls 2003) to *Political Liberalism* (Rawls 2005). *Political Liberalism* may be seen as a remedy to the problem of motivating people to endorse the principles of justice and to behave accordingly. Rawls says:

To understand the nature and extent of the differences [between Political Liberalism and A Theory of Justice] one must see them as arising from trying to resolve a serious problem internal to justice as fairness, namely from the fact that the account of stability in Part III of Theory is not consistent with the view as a whole. I believe all differences are consequences of removing that inconsistency (Rawls 2005, xv-xvi).

In the III Part of *Theory* the problem of motivation is faced and apparently resolved by adopting the argument of congruence. Rawls believes that people may sustain justice insofar as justice is part of

their good. Specifically, Rawls assumes that people in a fair society have an independent sense of justice and aspire to do what is right for its own sake. This does not mean that justice is always the best thing for them to do; indeed it may require of them to give up any of their specific purposes¹. Thus, the question is what kind of assurance one does have that it is realistically possible for people to affirm justice as fairness as part of their good. The congruence argument purports to show that under ideal conditions the judgements that would be made from the private perspective (judgements regarding a person's good) and the public ones (judgements of justice) will cohere. In sustaining this position Rawls refers to a complicated argument grounding in the Aristotelian principle: the main idea here is that the sense of justice is among human higher capacities and involves an ability to understand, apply and act on and from requirements of justice (Rawls 2003, 372-379).

I am insisting on motivation as motivation for justice is an indispensable condition of stability: a social order is hardly stable if people incline to pursue their particular good in place of the right for all. To be motivated for justice is the first assurance of stability. Stability is the central issue of the Rawlsian theory: it is here meant as an indicator of the feasibility of any political theory. "We checked to see [...] if justice as fairness is a feasible conception [...]; this forced us to raise the question of stability" (Rawls 2003, p. 508). A theory of justice should be feasible to be desirable: a theory that cannot be applicable is undesirable². A conception of justice is stable when its realization by the social system tends to bring about the corresponding sense of justice that induces citizens "to develop a desire to act in accordance with these principles and to do their part in institutions that exemplify them" (Rawls 2003, 119).

In *Political Liberalism* the issue of feasibility becomes more relevant. Rawls gradually became dissatisfied with the account of the just society as depicted in *Theory*. He finds the just society as designed in *Theory* to be *unrealistic*: it is quite unrealistic to maintain that people have the same reasons to support it and that they similarly view their good as congruent with the right. When the veil of ignorance is raised, citizens may realize that justice does not correspond to their ideas of good. That means to recognize the "fact of reasonable pluralism": under "the political and social conditions secured by basic rights and liberties of free institutions, a diversity of conflicting and irreconcilable – and what's more reasonable – comprehensive doctrines will come about and persist" (Rawls 2005, 36). He adds that pluralism "is not a mere historical condition that may soon pass away; it is a permanent feature of the public culture of democracy" (*Ibid.*). This clarification may be put in other words: pluralism is not an *external* constraint but it is a *fact about* liberalism and part of any ideal society. Liberal democracies promote pluralism as a consequence of reasoning under their institutions³.

On the premise of the fact of pluralism, the stability of the theory is now evaluated with regard to its appropriateness in constituting the focus of an overlapping consensus among different and conflicting comprehensive doctrines endorsed by reasonable citizens. For this stability to be realistically possible all the reasonable comprehensive doctrines affirmed by reasonable citizens must accept and endorse its political conception of justice in an overlapping consensus. When such an overlapping consensus exists, then all free and equal citizens endorsing reasonable comprehensive doctrines agree on the political conception of justice on the basis of their own particular comprehensive reasons. The matter of motivation is here resolved by assuming a freestanding political conception of justice that citizens may share. They come to defend this political conception as a common moral ground they can agree upon in spite of the diversity of their moral justifications⁴. It means that citizens may adhere to the

1 See Weithman (2010, ch. 7).

2 I may just mention the huge problem concerning ideal theory and its role to guide actions. On this point see Pasquali (2009) and Zuolo (2012).

3 On this point see Quong (2011).

4 By common moral ground Rawls means a political morality, not a comprehensive one, that is, a set of political moral values which may be subscribed by several comprehensive doctrines. See Rawls (1995).

principles of justice backed by their specific reasons: these principles may be adopted as a module that fits into and can be supported by various reasonable comprehensive doctrines that endure in the society regulated by it (Rawls 2005, 12-13). The principles of justice are worked up from certain “fundamental intuitive ideas” that are implicit in the public culture of a democratic society (Rawls 2005, 192). That explains why people may abide by them on the basis of different and even divergent views. They may be said however to converge on justice for “right reasons” (i.e. moral reasons), not for prudential reasons like opportunity or fear. Rawls stresses this point by declaring that the overlapping consensus on justice must not be confused with a *modus vivendi*. The *modus vivendi* is exactly a precarious equilibrium of forces that depends on “circumstances remaining such as not to upset the fortunate convergence of interests” (Rawls 2005, 147). Thus, the question of stability cannot be addressed by a *modus vivendi*; it can only be addressed by striving to realize an overlapping consensus on moral political values such that “stability is secured by sufficient motivation of the appropriate kind acquired under just institutions” (Rawls 2005, 142-143).

To sum up: in *Political Liberalism* Rawls starts by acknowledging the fact of pluralism as the result of living under liberal institutions. Pluralism is indeed the consequence of the exercise of the practical reason under free conditions. Given the fact of pluralism, the moral justification for justice as held in *Theory* turns out to be unrealistic: it is unrealistic that people agree on the same justification of justice. Instead, *justice as fairness* defended in *Theory* may be sustained by each person as a political conception, independently from their specific reason for endorsing it. The following point is how people may converge on this political conception: convergence may be assured once they adopt a *method of avoidance* according to which divisive issues are systematically excluded from the political agenda (Rawls 1985). If the project of justification in *Political Liberalism* is similar to *Theory*, now Rawls avoids relying on controversial moral positions and counts instead on ideas and reasons that are widely shared in democratic culture: the method of avoidance means to avoid claiming to truth to defend one’s own political position. Truth is irreconcilably divisive for Rawls: truth and politics must part company if justice (the theory and the practice of justice) wants to gain stability. Indeed, the so-called “political turn” consists in justifying justice by renouncing to defend one’s own moral justification as both exclusive and exclusively true⁵. Truth is to be replaced by reasonableness when political constitutionals are at stake. Reasonableness and not truth is the political standard of correctness: the objectivity of judgements about justice is characterized without reference to the notion of truth (Rawls 2005, 127). Lastly, avoidance of truth is not to be conflated with scepticism: to avoid claiming to one’s truth does not mean to deny it. It means only that we are allowed to use our deepest convictions in the construction of the political conception within certain limits, being those limits based on what people could reasonably accept (see Maffettone 2010, 269).

2. **Reasonableness and Unreasonableness** Let me now articulate the meanings of reasonableness⁶. So far I mentioned the notion of reasonableness as if it were a sort of “device” to connect comprehensive doctrines and the question of stability through the idea of an overlapping consensus. Such an overlapping consensus among comprehensive doctrines may be achieved only among that class of comprehensive doctrines that are reasonable, meaning that reasonableness qualifies the group of comprehensive doctrines that can find common ground within a liberal-democratic regime. In so doing, these doctrines may coexist albeit they diverge on truth. Then, reasonableness instead of truth may respond to the central question of *Political Liberalism*: “how is it possible for there to exist over time a just and stable society of free and equal citizens who still remain profoundly divided by reasonable religious, philosophical and moral doctrines?” (Rawls 2005, 47). Reasonableness responds to this

5 On the “political turn” in the Rawlsian thought about justice see Weithman (2010).

6 For a larger discussion on this point see Sala (2013).

initial problem of reconciling stability with pluralism: the stability of a fair society is secured by citizens' ability to mediate between their competing comprehensive claims. This citizens' ability is rooted in their reasonableness.

I now add further remarks on reasonableness. I firstly remind that according to Rawls a person is reasonable when she is willing to propose and honour fair terms of cooperation (Rawls 2005, xlii). Reasonableness indicates reasonable persons' ethical-political ability to share public reasons in a regime of reciprocity. Reasonableness may be here defined as the political virtue of citizens who show to be reasonable as they take others to be politically free and equal, and equally deserving of fair terms of social cooperation. Secondly, to be reasonable implies also to acknowledge the fact that, when freedom of thought and conscience is granted, people adopt reasonable yet incompatible comprehensive doctrines. In other words: to be reasonable involves to recognize the fact of reasonable pluralism. People are reasonable insofar as they support freedom of thought and conscience and recognize the consequence of this freedom in terms of reasonable pluralism⁷.

In summary: reasonableness implies two aspects – the willingness to propose and honour fair terms of cooperation and the willingness to recognize the fact of pluralism as the result of the free exercise of practical reasons – that substantiate what being reasonable means.

Having said that, reasonableness explains also how and why citizens come to distance themselves from their comprehensive doctrines: as they are reasonable, they may avoid to assert the truth of their beliefs and try to reach an overlapping consensus on liberal values. This capacity of avoidance plays as the evidence of their being reasonable. I said that reasonableness is a political virtue: it is in the political domain that reasonableness takes the place of truth. Indeed, the notion of reasonableness is strictly political: that is, it does not have to do with the contents of beliefs and with philosophical, moral or religious positions. On the contrary, it does have to do only with the attitude to tolerate others' positions as they are acknowledged as legitimate albeit different from ones' own positions. I conclude by saying that reasonableness is somehow the winner of a virtual game among values: it is *the* common value, it may be called the *common currency* of a society that vindicates its fairness.

Having recapitulated the meanings of reasonableness, what about those who do not “subscribe” to it? It seems a platitude to state that people who do not share the virtue of reasonableness are “unreasonable”. Rather, what Rawls means by “unreasonable” can be equated with planning to engage in cooperative schemes for the “wrong reasons” (Rawls 2005, 55). Unreasonable people are unwilling to honour, or even to propose, any general principles for specifying fair terms of cooperation. They may adhere to order not being properly motivated to support institutions: any adherence to them equates to a *modus vivendi* given that the reasons for adherence are prudential and not moral reasons. For Rawls, those people are to be kept under control as they would impose their beliefs on others: they are considered unreasonable not because their belief is thought to be false, but because they are ready to impose it on others⁸. In such cases, says Rawls, “the problem is to contain them so that they do not undermine the unity and justice of society” (Rawls 2005, xvi-xvii). Reasonable institutions have “the practical task of containing them – like war and disease – so that they do not overturn political justice” (Rawls 2005, 64, n. 19). Notwithstanding the fact that unreasonable people are a *fact* of any liberal democracy, Rawls trusts liberal institutions in their capacity to educate unreasonable people to cultivate in themselves a sense of justice. The fair society has a kind of educational or transformative role: the “unreasonable” are likely to become reasonable, because of the strong positive effects of living under liberal institutions.

7 On reasonableness in Rawls: Boettcher (2004); Mandle (1999); Archard (2001); Rasmussen (2004).

8 On this point see Waldron (2003).

Rawls hopes that in this way political stability increases and that the benefits assured by stability will eventually motivate the “unreasonable” to become reasonable (Rawls 2005, 163-168).

3. **Unreasonableness
Revisited**

Under actual circumstances, Rawls' hope may be misplaced. There may well be persons who will not adopt public reasons to defend their positions in the political debate. However, this does not imply that they are ready to impose their beliefs on others being unreasonable in the Rawlsian terms. The fact that they do not share values as freedom and equality in which the fair cooperation is grounded is not a sufficient reason to say that they will necessarily try to violate the terms of cooperation itself. I think instead that there may be more than one way to be unreasonable. Besides those unreasonable persons who actually represent a risk to stability, there may be people who are neither reasonable nor unreasonable. I call them “non-reasonable”⁹. Although they do not endorse the common values that form the overlapping consensus, they are ready to support liberal institutions, albeit for reasons that are not the “right ones”. Non-reasonable persons may support the liberal democratic order for *their own* reasons, which fall outside the domain of public reason and that may not be part of the overlapping consensus.

Unlike Rawls, I believe that nothing wrong happens if liberal institutions are endorsed by people divided at the level of their fundamental values. Stability is not necessarily threatened by the fact that institutions may be supported also by people who are unable to find any “continuity” between *their* values and the common values, between their beliefs and the civic value of reasonableness, but who are nonetheless ready to engage in cooperation with the others. Such a cooperative attitude, however motivated, may provide a strong enough motive to include such non-reasonable people in the liberal-democratic citizenry and treat them as political equals. For example, non-reasonable people of this kind may recognize liberal democratic society as the place to accomplish their mission on earth. Although their reasons for cooperation are unlikely to be translated into public reasons, they may equally count in favour of their sincere participation in a fair system of social cooperation. This implies also to contradict Rawlsian thesis about the connection between reasonable doctrines and reasonableness as an attitude: people may behave “not unreasonably”, despite the fact that they hold unreasonable doctrines (in the Rawlsian sense) and are not ready to renounce them for any reason¹⁰.

My conclusive point is that actual circumstances of politics should not be neglected by appealing to an ideal overlapping consensus, by which common moral values should be shared. Non-reasonable people may in fact participate in the public debate by putting forward reasons that are neither public nor shared by other reasonable citizens. Public debate may be opened also to those who may offer *their* reasons that are hardly public reasons: their reasons matter from their specific point of view. The fact is that in realistically plural societies several groups advocate policies for their non-sharable personal reasons. An enlarged public debate is not expected to lead necessarily to an overlapping consensus on the basic liberal values. Instead, it allows all citizens, including the “non-reasonable”, to start from their respective points of view and exhibit *their* reasons publicly (Archard 2001). These “non-reasonable” people may thus be included in an enlarged public debate, instead of being deprived of such an opportunity as it happens in the Rawlsian society. Indeed, not all citizens are reasonable in the Rawlsian sense and the problem of how to deal with the ‘non-reasonable’ is not a mere accidental matter, but is a crucial fact of politics with which the liberal theories of justice should concern themselves.

⁹ For a more detailed assessment of non-reasonableness as alternative to both reasonableness and unreasonableness in the Rawlsian terms, see Sala (2013).

¹⁰ Rawls inclines to see as reasonable those people who hold reasonable doctrines. By contrast, not all people who hold unreasonable beliefs should be called unreasonable. See Rawls (1997). About the relationship between unreasonable doctrines and reasonable behaviour see the exchange among Kelly and McPherson (2001) and Quong (2004). A comment on it in Sala (2013).

What I mean is that public justification may be conceived of as a wider practice in which the “non-reasonable” may participate by putting forward their special reasons that are unlikely to be acceptable to all. My suggestion is that although the reasons of the non-reasonable are likely to be unacceptable to reasonable people, they may nonetheless be made intelligible to them, that is, they may be exchanged in public debate (Gaus 2010). Thus, the reasons of the non-reasonable people should not be dismissed from public debate in advance of an argument simply on the grounds that reasonable people disagree with them. If a political theory aspires to be realistic it is crucial that it takes into account the actual conditions in which persons may have both moral and non-moral reasons to accept the liberal-democratic order notwithstanding the fact that they do not share its essential values¹¹.

¹¹ On realism in political theory with regard to Rawlsian tradition: Galston (2010), Mason (2010), Horton (2010), McCabe (2010).

REFERENCES

- Archard, D. (2001), "Political Disagreement, Legitimacy and Civility", *Philosophical Explorations*, 3, 2001, pp. 201-222;
- Boettcher, J. W. (2004), "What is Reasonableness?", *Philosophy & Social Criticism*, 30, 5-6, 2004, pp. 597-621;
- Galston, W. (2010), "Realism in Political Theory", *European Journal of Political Theory*, 9, 4, 2010, pp. 385-411;
- Gerald, G. (2010), *The Order of Public Reasons*, Cambridge University Press, Cambridge, 2010;
- Horton, J. (2010), "Realism, Liberal Moralism and Political Theory of Modus Vivendi", *European Journal of Political Theory*, 9, 4, 2010, pp. 431-448;
- Kelly, E. and McPherson, L. (2001), "On Tolerating the Unreasonable", *Journal of Political Philosophy*, 9, 1, 2001, pp. 38-55;
- McCabe, D. (2010), *Modus Vivendi Liberalism. Theory and Practice*, Cambridge University Press, Cambridge, 2010;
- Maffettone, S. (2010), *Rawls. An Introduction, Polity*, Cambridge, 2010;
- Mandle, J. (1999), "The Reasonable in Justice as Fairness", *Canadian Journal of Philosophy*, 29, 1, 1999, pp. 75-108;
- Mason, A. (2010), "Rawlsian Theory and the Circumstances of Politics", *Political Theory*, 38, 5, 2010, pp. 658-683;
- Pasquali, F. (2009), "Rawls's Realistic Utopianism: a Critical Discussion", WP-LPF 5, 2009, pp. 1-38;
- Quong, J. (2004), "The Rights of Unreasonable Citizens", *Journal of Political Philosophy*, 12, 3, 2004, pp. 314-335;
- Quong, J. (2011), *Liberalism without Perfection*, Oxford University Press, Oxford, 2011;
- Rasmussen, T.W. (2004), "Defending Reasonability. The Centrality of Reasonability in the Later Rawls", *Philosophy & Social Criticism*, 30, 5-6, 2004, pp. 525-540;
- Rawls, J. (1985), *Justice as Fairness: Political Not Metaphysical*, *Philosophy & Public Affairs*, 14, 1985, pp. 223-252;
- Rawls, J. (1995), "Political Liberalism: Reply to Habermas", *Journal of Philosophy*, 92, 3, 1995, pp. 132-180;
- Rawls, J. (1997), "The Idea of Public Reason Revisited", *University of Chicago Law Review*, 64, 3, 1997, pp. 765-807;
- Rawls, J. (2003), *A Theory of Justice* [1971], The Belknap Press of Harvard University Press, Cambridge, 2003;
- Rawls, J. (2005), *Political Liberalism* [1993], Columbia University Press, New York, 2005;
- Sala, R. (2013), "The Place of the Unreasonable People Beyond Rawls", *European Journal of Political Theory*, forthcoming;
- Waldron, J. (2003), "Toleration and Reasonableness", in C. McKinnon and D. (eds.), *The Culture of Toleration in Diverse Societies. Reasonable Tolerance*, Manchester University Press, Manchester and New York, 2003, pp. 13-37;
- Weithman, P. (2010), *Why Political Liberalism? On John Rawls's Political Turn*, Oxford University Press, Oxford, 2010;
- Zuolo, F. (2012), "Being Realistic without Realism. Feasibility and Efficacy in Normative Political Theories", LPF-WP, n. 3, 2012, pp. 1-29.

SESSION

4

SESSION 4

SPECIAL ISSUE

*Emanuele Caminada (Universität zu Köln),
Barbara Malvestiti (Università degli Studi di Milano)*
An Interview with Martha Craven Nussbaum:
Political Emotions. Why Love Matters for Justice

*Valentina Bambini, Cristiano Chesì, Andrea Moro
(Istituto Universitario di Studi Superiori di Pavia)*
A Conversation with Noam Chomsky: New Insights on Old Foundations

EMANUELE CAMINADA

Universität zu Köln

emanuele.caminada@gmail.com

BARBARA MALVESTITI

Università degli Studi di Milano

barbaramalvestiti@virgilio.it

AN INTERVIEW WITH
MARTHA CRAVEN NUSSBAUM:
POLITICAL EMOTIONS.
WHY LOVE MATTERS FOR JUSTICE

Martha Craven Nussbaum (<http://www.law.uchicago.edu/faculty/nussbaum/>) is one of the most distinguished philosophers of our time. If the starting point of her practical philosophy is the Aristotelian theory of the good life, the study of emotions in ethics is rooted in her profound knowledge of ancient philosophy. She takes part in global philosophical debates about liberalism, feminism, multicultural political philosophy, and international justice. Along with Amartya Sen, she defends the Capability Approach in development policy. Currently, she is researching on the meaning of religion and its role in society.

As Albertus Magnus Professor 2012 (<http://amp.phil-fak.uni-koeln.de/>) Martha Craven Nussbaum gave three lectures (June 19th-21st) about her current research project, disclosing some of the topics of her forthcoming *Political Emotions: Why Love Matters for Justice*, Harvard University Press (2013).

Intrigued by her inspiring talks we ask her more about her philosophical project.

1) *The tradition of phenomenology attests that sentiments, feelings and emotions are intentionally related to values, which are the cognitive content of emotional life. How do you think that emotions, values and norms are related?*

In my view emotions always contains appraisals of an object as either good or bad for the creature who has the emotion. So in that sense they contain values. But the values might not be at all reflective, and they might or might not be norms, since I think by “norms” we mean a set of social agreements about value. When a rat flees a predator, there is emotion, and in that minimal sense there is value (death is bad), but no norms, since to have norms you have to be capable of abstract reflection to at least some degree.

2) *Does a distinction between “needs of soul” and pure desires exist?*

You’d have to explain this one to me. I am not sure what “needs of soul” means. If you mean the contrast between occurrent desires and deep persisting needs that express our sense of our identity, then yes, there is of course such a distinction. I might desire an ice cream cone, but I don’t need it. I might or might not feel an occurrent desire for freedom of speech, but I need it, in the sense that its

presence is a necessary condition for a life worthy of my human dignity.

3) In your current research project you aim to connect both pillars of your philosophy: your cognitive theory of emotional life and your capability approach as a theory of justice. Why are you convinced that social philosophy needs both of them and why are they deeply interrelated?

Obviously society needs a map of where it is going. I think that a decent society will frame that map by thinking about justice, and a theory of justice is one very important part of framing social goals well. But to achieve those goals, and to keep them stably once we have them, we need to think about the emotions. Some emotions help society achieve its goals and others do not, and we'd better understand these connections as well as we can.

4) What do you think about the relationship between a right and the correlative duty? Which one exists first? Are they co-originary? If a man was alone on an Island, would he have rights? And duties?

I think rights and duties are correlative, and that they both exist independently of political society. A solitary person has both rights and duties, but there is no institutional structure to guarantee the former or to channel the latter (as through a tax system, for example).

5) Is the existence of good reasons for rights also a necessary, even though not sufficient, condition of their effectiveness?

Neither. We might have an effective rights regime for which nobody could articulate good reasons, although that is likely to be a rare occurrence. And there have always been good reasons in favor of many rights that are not effective (such as the equal rights of women, or people with disabilities).

6) Does a right exist independently of legal guarantees?

Yes, since rights inhere in human dignity itself.

7) What basic values are presupposed in your political liberalism? Do you have a hierarchy of values? How are those values related to personal capabilities?

I have identified, provisionally, ten Central Capabilities that a good society ought to protect up to some suitable threshold level. They are not hierarchically ranked: all are regarded as necessary for a life worthy of human dignity. All are personal capabilities, that is, opportunities of persons. But though they are not ranked, we might in particular circumstances judge that some take priority, because they are the best intervention points in that context to promote, over time, the realization of the whole list. Those are what Jonathan Wolff and Avner De-Shalit call "fertile" capabilities: the ones that engender others.

8) What reasons bring you from a kind of social-democratic political philosophy to a liberal-egalitarian one? Are you committed to a kind of social ontology or metaphysics concerning the relationship individual-society?

I don't understand this. My view is both social-democratic and liberal (meaning allocating a large place to fundamental liberties), and it has not changed in those respects. And political liberalism of the Rawlsian sort, which I endorse, holds that one should not ground any political principles in metaphysics or ontology: the principles should be justified by "freestanding" ethical argument. That is what I hope I have done.

9) Referring to your political principles you underlined that they have to be ethically motivated and not metaphysically. Are you claiming that your political principles are metaphysically neutral or that you have to bracket them in order to bring these results of your philosophical thought in the political debate?

I mean that they are thin, not grounded in metaphysics of the type that divides people along lines of their comprehensive doctrine. People can then interpret them metaphysically if they want, or not if they don't want. Take the idea of human dignity: it is a resonant ethical idea, but it takes no stand on the question whether human beings have immortal souls. Someone who believes in the soul can always understand dignity in connection with that idea, but someone who doesn't believe in the soul is not forced to accept that doctrine as a condition of signing onto the political principles.

10) In the historical part of your working project you refer to a cosmopolitan tradition: Rousseau, Herder, Mazzini. Particularly you defined Mazzini's European patriotism as a form of pietism that tries to let people transcend their self-interest toward the common good. Can you say more about the way you link Rousseau, Herder and Mazzini's republicanism and how do you move further in your political thought?

Actually I just took 500 pages to answer this question, and I can't imagine any way of doing so briefly here. For one thing, the two thinkers I eventually agree with most in this tradition are J. S. Mill and Rabindranath Tagore, and since Tagore's thought is virtually unknown in Europe it would take some time to spell it out. So you will have to wait for the book, I'm afraid.

11) An important issue of this project is to shape the idea of a kind of cosmopolitan civil religion, that should not be a rival to other religions or personal preferences: it has to be inclusive. How far away is the humanistic ideal from yours? To what extent is this ideal of civic religion related to material values that are culturally embedded or is there a kind of telos to a civic religion that is most neutral to historical ones?

Actually, I do not use the word religion, that word is used by the thinkers I discuss. I think it's misleading to use that word, except in scare quotes, because it suggests that agnostics and atheists are not included as equals.

12) You focused your attention on two central emotions in human life: compassion and disgust. Referencing animal psychological and cultural research, you claim that emotions are narrow and eudaimonistic: they are blind to interests that are not personally anchored. Therefore you claim that the affective education of the person is necessary in order to shape a decent society. Humans have to learn to extend their compassion to others and to reconcile themselves to their animality, because the power of disgust can be very aggressive in excluding abnormal subjects. Today, human dignity is a central topic in political and juridical debates. What do you think about this concept? What kind of relationship exists between human dignity and feelings of disgust and shame? And between human dignity and reserve?

Actually, my book does not focus only on compassion and disgust. It has large section on love of individuals, love of country, fear, shame, envy, and quite a few others. Nor do I hold that emotions are always narrow: I say that they begin that way, but they can be broadened by education. Dignity is a very important concept, but it does little argumentative work on its own, since it is so vague and so disputed. It should be understood as part of a family of political concepts and principles that should all be justified holistically. Moreover, we need to think about the dignity of non-human animals as well, and make sure that our concepts and principles show respect for that.

13) In the Greek tradition you discovered two paradigmatic genres of emotional education that relate to these fundamental human needs: the tragedy, that can arouse compassion and the comedy, that has the power to rehabilitate human bodily fragility and vivacity. Why do you think that both genres are fundamental to democratic societies? How can a modern liberal society follow these intuitions? In what extent do you believe that political liberalism needs a liberal education in arts and humanities? Which role does imagination play in political and juridical judgment?

This is another huge question, the answer to which is at least three books of mine ... But briefly: tragic

festivals cultivate extended compassion for human predicaments, comic festivals make us well-disposed to the body. We can pursue those goals through a variety of media: political rhetoric, public holidays, public music and art, memorials, parks, book clubs, and yet others.

14) You are looking for normative criteria in order to establish a “decent society”, looking for the bare minimum of chance and equality for everyone. This project, based on the philosophical principles of political liberalism, wants to shape human emotions in order to let these principles be stable in modern societies: it suggests a more Platonic than Aristotelian policy. Why is your project different to Plato’s Republic? Are you looking for the minimum decency or for the best possible societies?

My idea of a threshold is not an idea of a “bare minimum”. It is an idea of what is necessary for a life worthy of human dignity, and that is usually going to be quite ample. I don’t see why you call the project Platonic. Indeed Aristotle’s Rhetoric supplies essential guiding points. And since the heroes are Tagore and Mill, it is realized in the midst of a vigorous critical culture that protects spaces for dissent and individual experimentation. Plato would have hated almost everything I say.

15) You work out human and animal psychology: What kind of animal do you think we are? To what extent do you agree with Aristotele’s definition of man as “zoon politikon”?

I think we are animals who are social, which is what Aristotle means, who are capable of great love and also terrible hatred, hatred of one’s own kind unknown in other species. This ambivalence derives in part from our initial helplessness and our contrary expectation of being in control: the tension between those two aspects of our humanity leads to some very difficult problems in human life. There is no non-human animal who wants not to be an animal, who repudiates the very idea of animality and mortality, and this desire to control everything and transcend our finitude is a recipe for political difficulty.

16) Your case studies are taken from classic Greece, modern American history and contemporary India. In particular you refer to Gandhi and Tagore. What are the most relevant elements of Indian political thought? Do they differ from western ones in form and content?

I’m not sure why you leave out Nehru and Mill, who are my two greatest personal heroes and friends. But anyway, I don’t think you’d speak of “American political thought” or “German political thought”. What would that mean? What do Heidegger and Habermas have in common? Or Rawls and Nozick? When we know a country well, we realize that it isn’t a single thing and that little of interest can be said at that level of generality. Why should things be different with India? I choose Tagore, Nehru, and Gandhi because I find their ideas valuable and illuminating, not particularly because they are exemplars of “Indian thought,” whatever that would mean. And I also love B. R Ambedkar, who disagreed strongly with both Nehru and Gandhi. So one thing I value in India thought of the mid-twentieth century is its rich debates on matters of fundamental importance.

AN INTERVIEW WITH MARTHA CRAVEN NUSSBAUM:
POLITICAL EMOTIONS. WHY LOVE MATTERS FOR JUSTICE

SELECTED BIBLIOGRAPHY OF MARTHA CRAVEN NUSSBAUM:

The Therapy of Desire. Theory and Practice in Hellenistic Ethics, Princeton University Press, Princeton, 1994;
Women and Human Development. The Capabilities Approach, Cambridge University Press, Cambridge, 2000;
Upheavals of Thought. The Intelligence of Emotions, Cambridge University Press, Cambridge, 2001;
Hiding from Humanity. Disgust, Shame, and the Law, Princeton University Press, Princeton (N.J.)-Oxford, 2004;
Frontiers of Justice. Disability, Nationality, Species Membership, The Belknap Press of Harvard University Press, Cambridge Mass., 2006;
The Clash Within. Democracy, Religious Violence, and India's Future, The Belknap Press of Harvard University Press, Cambridge Mass., 2007;
From Disgust to Humanity. Sexual Orientation and Constitutional Law, Oxford University Press, Oxford, 2010;
Not For Profit. Why Democracy Needs the Humanities, Princeton University Press, Princeton, 2010;
Creating Capabilities. The Human Development Approach, The Belknap Press of Harvard University Press, Cambridge Mass., 2011.

VALENTINA BAMBINI

CRISTIANO CHESI

ANDREA MORO

Istituto Universitario di Studi Superiori di Pavia

valentina.bambini@iusspavia.it

cristiano.chesi@iusspavia.it

andrea.moro@iusspavia.it

A CONVERSATION WITH NOAM CHOMSKY: NEW INSIGHTS ON OLD FOUNDATIONS

“It is important to learn to be surprised by simple things – for example, by the fact that bodies fall down, not up, and that they fall at a certain rate; that if pushed, they move on a flat surface in a straight line, not a circle; and so on. The beginning of science is the recognition that the simplest phenomena of life raise quite serious problems: Why are they as they are, instead of some different way?”
(Chomsky 1988)

One of the leading intuition of Noam Chomsky’s research on language and cognition is that it is important to learn to be surprised by simple things. One can take the questions presented here as an example of this approach. The questions were proposed by undergraduate and graduate students, reorganized by us and presented to Noam Chomsky on occasion of his visit to the Institute of Advanced Study in Pavia on September 15th, 2012. Simply and spontaneously, the questions raised by the students clustered around a number of fundamental topics concerning the architecture of the language faculty, its relation with other systems. Starting from the mathematical properties of language, Chomsky answered students’ questions by discussing theoretical and epistemological consequences of the research on the biological foundation of language. He also posed new questions that such rational inquiry opens up and that can possibly get an answer in a future. The call for simplicity is reflected in the content of the answers Chomsky provided, and in the style as well. The discourse unfolds smoothly guiding the hearer – and the reader – through leading issues in the Chomskyan approach, from the foundation of the field (Chomsky 2004) to the ultimate synthesis (Chomsky 2011), across the milestones that made the scientific study of language possible in the modern era and still lucidly indicate where we should be heading to¹.

1) Can we study evolution of language? Is there any evidence that language evolved?

The subject “evolution of language” is a very fashionable one. In the last several decades there is a huge literature, libraries of books, international conferences, papers pouring out all the time. There are a few problems about it. One problem is that the topic doesn’t exist. Small problem... Furthermore, everyone knows it doesn’t exist. Evolution involves changes in the genomic characters of the organism. Languages are not organisms, they don’t have genomes, they don’t evolve. Languages change, but they don’t evolve. What evolves is the language capacity of users of language, i.e., human beings. So problem number 1 is: the topic doesn’t exist. Problem number 2 is that the work that has been done on the topic is utterly unlike anything that has been done in evolutionary biology. For

¹ This contribution contains the faithful transcript of Chomsky’s seminar. The notes provide some additional references. Although the contribution has been organized and revised by all authors, V. Bambini is responsible for questions 1-2, 8-11 and for the notes, while C. Chesì for questions 3-7. We would like to thank the students who participated to the initiative and submitted their questions, and the Collegio Ghislieri in Pavia for hosting the seminar.

example, suppose a biologist submitted a paper on the evolution of the eye. Consider that he has no idea of what an eye is and says that an eye is maybe something that you use to watch television. People would laugh. You couldn't submit a paper like that. But that is exactly what the literature on the evolution of language is about. It doesn't tell you what they think language is, just that language is something used for communication, which is about like saying that an eye is used to watch television. One of the many uses of language is communication, but that doesn't tell you anything. Other problems arise as soon as you begin to look at the work. For example, a lot of the work by really sophisticated people argues that frozen expressions in language are fossils from earlier periods. They tell you something about the earlier stages of language. It is usually English that is studied; so take a frozen expression in English. If it is a fossil, it is a fossil from a couple of hundred years ago, maybe a thousand years. What can that tell you about the history of language? Nothing. That is a tiny blink of an eye in the last phase of a long history of language.

The subject is one of the strangest subjects I have ever seen. Now, there is a subject: evolution of the capacity to use language, or to acquire language, the human language faculty. That is the real subject. But there is a problem with that subject too: we don't know anything about it. Actually, we know two facts about the evolution of the language capacity. One of the facts is sufficient to tell you that the study of the evolution of language is mostly a waste of time. For example, one of the things we know about the evolution of the language capacity is that nothing has happened for at least 50 thousand years since humans left Africa. And the evidence for this is pretty strong, in fact very compelling. If you take an infant from an Amazonian tribe in Brazil, a tribe that maybe hasn't had contact with other humans for maybe 20 thousand years, the child first of all learns Portuguese instantly, with no problems, and if the infant was brought to Pavia, he would be speaking just like you, exactly the same. And conversely, if you take an infant here and dump him in the Amazon, he would speak exactly like the children of that tribe. We don't know any deviation from this: not every case has been studied, but such a wide range of cases have been observed that it is almost certain that this is universal. What that tells you is that the capacity for language hasn't evolved in a way beyond historic times back to the first humans leaving Africa, or at least their ancestors, roughly 50 thousand years ago. So at least in that period there has been no evolution of the language capacity.

The second fact that we know with not full confidence but kind of plausibility is that, if you go back quite a short period before that – a short time in the evolutionary time, maybe 50 or 100 thousand years before that, which is nothing in evolutionary time, there is no evidence that language existed at all. There is substantial archeological evidence that somewhere in that very narrow window, roughly 75 thousand years ago, there was a sudden explosion of evidence of creative activity, symbolic representation, indications of the phases of the moon, complex social organization, and various rituals, all sort of things that indicate that something happened in a very small group, maybe a small hunter-gatherer tribe, maybe a hundred people, and then spread. So, within again a short period of time, there was Homo Sapiens: in fact that is the origin of cognitive Homo Sapiens. Physical, anatomical Homo Sapiens, goes hundreds of thousand years back beyond that. But something happened cognitively in one small group and essentially took over. You can guess what the date is: within the last hundred thousand years probably. That tells you that some event took place, brief change which provided our ancestors with the creative capacities that we all have, and we all have essentially identically, because we are all different descendants from that small group who among their other talents succeeded in a way beyond everybody else. The human species is kind of unusual and all related species have been wiped out. There is no comparative evidence. There were plenty of other hominids – we know that from archeological evidence – but they are all gone. Neanderthal lasted until maybe 30 thousand years ago (there is a very slight interbreeding with Homo Sapiens, mostly in Southern Italy, incidentally), but essentially everything else was wiped out. And in fact it goes beyond that. If you look at the spread of Hominids, going back a million years or longer than

that, as our ancestors spread, megafauna, i.e., big animals, disappeared. It was thought for a long time that it had to do with the expansion of the ice ages, but it turned out that, if you look at close dating, wherever proto-humans appeared everything else disappeared. We are now in the process of finishing it off... The end result is that there is essentially no comparative evidence. There is evidence that something happened within a very short period, essentially suddenly from an evolutionary point of view, and nothing has changed since. And that is the evidence. There is essentially no other evidence about the evolution of language².

If you take a look at the literature, what is studied is a different topic: language change. Undoubtedly languages changed. We are not talking Latin, we are not talking Sanskrit, proto Indo-European or whatever came before that. Undoubtedly languages changed, but that is not evolution. If you want to study language change, that is a serious discipline. Historical linguistics is a serious discipline, it has rigorous standards: you can't say "maybe this happened, maybe that happened". That doesn't count in historical linguistics. But that is what the work is in so-called evolution of language; it is what evolutionary biologists sometime call "just so stories": "maybe this happened, maybe something else happened". That is not serious work. So, if there is something to learn, as I think there is, we can ask what small change could have taken place that could have given rise to something like the language faculty. There is linguistic work which I think shed some light on that. I won't go into it. But we have to look for some small mutation that caused some rewiring of the brain that provided the essential properties of human language which don't exist anywhere else. And I think there is work on it. Beyond that, to talk about the evolution of language is just cutting down forests for no purposes, as far as I can see.

2) Is there any new contribution from neuroscience to the understanding of language?

Yes, there is. In fact the best work I know is Andrea Moro's work. There is a fair amount of interesting work. But there is one result that is quite far-reaching and that is what has been done here³. Cutting it down to its simplest form, it is something like this: taking speakers of some language, German let's say, and presenting them with two kinds of non-sense languages, languages they don't understand. One of them is modeled on say Italian: it has the properties of Italian, but non-sense words. The other is a language designed to violate what appear to be universal properties of language. The most interesting case that was studied, the one with the most far-reaching results, has to do with the linear order. For example, in Italian, if you want to negate a sentence, there is a negative particle which appears in a certain position in the sentence and it has to appear in that position, a fairly complex position: the position is defined by various structural relationships. You can make up a non-sense language in which negation is much simpler: you put it in a particular linear order, maybe say the third word of the sentence. If you want to negate a sentence, you take the sentence and put the negation particle in the third word. Computationally, that is a lot simpler. But no human language works like that. Every human language puts the negative particle in some structurally defined position, and this incidentally generalizes. Linear distance counting doesn't seem to exist in human language. There is nothing like third position or find the shortest, the closest word. I will give you a simple example. Take interpretation of adverbs in complex sentences: you can show it also in brief sentences. Take the sentence "Eagles that swim fly". Put the word "instinctively" in front of it: "Instinctively eagles that swim fly". Everyone knows, every young child knows that "instinctively" goes with "fly", it doesn't go with "swim". It doesn't go with the closest verb. If it is "Instinctively eagles that fly swim", you don't understand "instinctively" to go with "fly", although that is the only thing that makes any sense: you understand "instinctively" to go with "swim". "Instinctively" is not finding the closest verb: it is finding a remote one and one which happens to be the structurally closest by an abstract notion of

² For a paleoanthropological perspective, see Tattersall (1998).

³ For a comprehensive presentation of this and other experiments on possible and impossible languages, see Moro (2008).

distance. And that holds for every linguistic construction, every language. Languages don't use simple computational techniques like closest or third. They use complex computational techniques involving structural distance and structural position. Again, universally. Now, going back to the language modeled on Italian and the invented language with the simple position for negation, say third position, Moro and colleagues' experiment found that normal speakers could solve the language modeled on Italian: they learned it very quickly. As for the language that uses the simple computational principle, they could solve it, but it was a puzzle. So, different areas of the brain were activated, not the normal language areas. There is comparable work with aphasics, actually with idiot savant cases, cases of a person who has tremendous, fantastic, language abilities, but very limited cognitive abilities. Neil Smith has done work on it⁴. Somewhat similar studies that are behavioral and not neurolinguistic found essentially the same as what Andrea Moro found with the investigation of the activation of brain areas. The normals could solve the invented language problems that violated Universal Grammar, but in a complicated way: they treated it like any other puzzle. The idiot savants, the one with limited cognitive abilities but excellent language abilities, couldn't solve it at all. They can't solve puzzles. This evidence from neuroscience compared with evidence from behavioral studies provides independent and very significant evidence that languages just don't use simple computational procedures. They don't use linear order, linear distance, fixed positions. That tells you something about the language faculty. That is the kind of neurolinguistic work which can be very significant, I think. A lot turns on the question of whether languages use linear order: first semantics, first syntax, do they have rules that involve linear order, linear position? There is a lot of linguistic consequences to that choice. Neurolinguistic evidence conforms to other evidence in indicating that, if you want to understand the human language faculty, you have to ask what property is it that would lead automatically to use of structural position and structural distance, while ignoring linear position and linear distance. And there are interesting results about that, I think.

That relates very closely to the question of what small changes would have taken place that would have led to the sudden appearance of language in the first place. For those of you that know the linguistic literature, what is suggested in both cases is that the change that took place was the sudden emergence of the simplest combinatorial operation that is unbounded in scope. It is an operation that takes two entities already constructed and forms out of them the simplest possible new entity, the simplest possible will not involve linear order, because that is more complicated, it won't involve any changes in the two entities put together, and in fact what it will form is just the set of the two, and that turned out to give you the basis. An operation like that is embedded somehow in every computational system: it is the simplest one possible and it does give you an unbounded array of hierarchical structure, and any property of the system that is emerging will depend on structural and not linear distance, structural and not linear position⁵. That has plenty of consequences.

3) What is the role of the notion of simplicity in modern formal linguistics, i.e., in technical terms, Minimalism? Is it the end or the starting point?

Simplicity is simply the core notion of science. That goes back to Galileo. He argued that nature is simple and it is the task of the scientist to show it. When you look at phenomena, they look extremely complex and diverse, and the task of the scientist is to show that this is a superficial misunderstanding: if you look more deeply, you will somehow figure out that there are simple elegant rules and principles. The whole of science is based on that. If you do not do that, you are not doing science, you are collecting data, you are flower collecting, which is ok, but this is not science. Furthermore, there is a substantial philosophical literature in this respects (Goodman⁶ and others) which shows that the search for simplicity is identical with the search for explanation, and you can

4 On idiot savants, see Smith & Tsimpli (1995), where the case of Christopher is described.

5 This operation is technically called "merge" (Chomsky 1995).

6 The reference is to the philosopher Nelson Goodman, broadly known for the "grue and bleen paradox", introduced to highlight certain problems of induction (Goodman 1954).

see why, even superficially: the more complex your account the less deep the explanations. If you can make your account simpler, your explanations are deeper, and the whole purpose of rational inquiry is to find explanations, science as well. That is the driving force behind the study of language as well, and what is called “Minimalism” – maybe the name shouldn’t have been given to it – is just the latest stage in the effort to try to find simpler explanations⁷. It is a seamless continuation of everything that has been done in the study of language at least for the last 60 years, since the modern study took shape. In the early days, 60 years ago, it looked as if languages were extremely complex and very diverse: almost anything you can imagine could happen. Over the years it has been gradually found that, if you look at it properly, diversity is much more constrained and there are very limited options for change, and the things that look very complex on the surface often have deeper explanations that work. The case that I mentioned is just a quite interesting one: the fact that simple computational procedures are not used but rather things like structural distance and structural position. That tells you a lot: one of the things that it tells you is that the core operation in language is the one I mentioned, the simplest computational operation which could have emerged from some simple mutation, but here we are getting into questions about genetics and brain science that are not understood⁸. But it is quite possible that some very simple mutation could have rewired the brain to provide the simplest computational operation, which is unknown in the organic world, since it appears in the language and nowhere else.

4) Doesn't memory structure constrain grammar as the sensory-motor and the conceptual-intentional external systems do?

Memory structure certainly constrains usage, in fact humans have quite limited memory as compared with a number of other organisms, for example certain birds like crows. A crow can remember where it has hidden 10 or 20 thousand seeds, and not only where it has hidden them, but what quality the seed was, so that it will go first after the seeds that are going to deteriorate more quickly. Furthermore, the crow pays attention to other animals around that might see it hiding the seed and, if it sees that there is another crow around, it will wait that it goes away and re-bury the seed somewhere else. These are feats of memory that are absolutely unconceivable for humans: humans can get up to about 7 not 20 thousand: that is an order of magnitude difference. The narrowness of memory is so that the number 7 is almost uniform across organisms (crows happen to be unusual): 7 plus minus 2 is a famous formula in the study of short-term memory that goes back to George Miller, 50 years ago⁹. Humans are in that range, a normal range, and that does constrain things you can say: it constrains numbers that you can add in your head, for examples. But that tells you nothing about language, just as it tells you nothing about your knowledge of arithmetic. Your knowledge of arithmetic enables you to add numbers of arbitrary size, and you can demonstrate that: if you give a person more memory and more time (say paper and pencil), there is no limit on the size of the numbers you can add. That tells you that you have an internal procedure, technically it is called the “generative procedure”, which assigns a sum to any two arbitrary numbers. But memory constraints and time constraints limit how far you can go in real-time. You can go on and on indefinitely, if you have more and more time. Language is exactly the same. There is no indication that memory plays any role in the nature of language any more than it does in the nature of arithmetical knowledge. But of course, it plays a role in your use of language, just like in your use of your ability to do arithmetic, or any other generative process. We have to distinguish between the nature of the system and the use of the system: this is a very fundamental difference. In the study of language, this is often called the difference between

⁷ On the origin of the term Minimalism, which aims at reducing linguistic levels to the minimum number required by virtual conceptual necessity, see Moro (1996).

⁸ For an extensive discussion on computational operations, see De Palma (1974) and Chesi (2012).

⁹ See Miller (1956).

“competence” and “performance”: your internal knowledge, i.e., the internal structure, and what you do. It is considered a controversial distinction, but this has to be a confusion, since it is a conceptual distinction that cannot be avoided. There is a comparable distinction that is so obvious that it is just taken for granted everywhere else in biology. Suppose someone is studying the digestive system and doesn't know exactly how the digestive system works, i.e., the nature of the system: you don't just look at the performance, like if the person has intestinal flu or if he just ate a big meal; you don't pay attention to that, you abstract away from it, if you want to understand the nature of the system. That is so obvious that it is never mentioned: it is just taken for granted in all inquiry. But when you talk about language, for some reasons, rational assumptions dissolve very quickly. It is typical when we talk about ourselves: it can be very hard to be rational. But if you use normal rational standards, the distinction is obvious and it tells you where memory plays a role as far as we know: not in the organization of the system but in the use of the system, of course.

5) What do you think is the most critical computational aspect of language for a computer program that aims at understanding it?

First of all, no computer program should aim at understanding language, because that is utterly hopeless. It is much too elevated goal. We don't even know what understanding language means. You can't try to construct a program that does something that you can't characterize: that is not a feasible task. Understanding of language involves some many complex things that we have no grasp of: it is not a formulable task. You can ask narrower questions. For example, you can construct a computer program that will determine from the sequence of sounds or some other representations the internal structure of the object that you construct in your mind that gets interpreted in ways that we don't understand. That you can do: it is called a “parsing program”, technically, and it raises interesting questions. Just to go back to the question of linear order that I have mentioned, a parsing program would be much simpler if it could use things like linear order: say in the sentence “Instinctively eagles that fly swim”, a program that will tell you that “instinctively” goes with “fly” will be far simpler than the one that we use that tells you that it goes with “swim”. But language doesn't offer this possibility; in fact there is a striking conflict – we find over and over – between language design and ease of use of language. They are in conflict. Language is designed in ways that increase the difficulty of use. Put it differently, communicative efficiency and computational efficiency conflict in many cases, and those are quite interesting cases: in every single case that is known computational efficiency wins and communicative efficiency loses. So whatever language is about, it doesn't really care much about communication: it cares about internal elegance, being a simple system. And if you think about how language must have evolved, very recently, that makes sense: it was an almost instantaneous emergence. Whatever happened at that interesting moment, 75 thousand years ago, under no external pressure, no selectional pressure – it just happened in somebody's brain, it would naturally take the simplest possible form. Using an analogy that I have used occasionally and that Andrea has used, it should become like a snowflake: it just takes the simplest form that the proprieties of physics forces it to take. That is what seems to have happened: in so far as we understand anything about language, it seems to take the simplest form and not care much about the consequences for use, in particular for communication. There is a lot of evidence pointing to that, and it has many interesting consequences. To go back to the literature on evolution of language, it bears on that too: standard view is that language evolved as a means of communication, but the evidence is overwhelmingly against that. It seems to have developed in a way that harms communication and doesn't care about communication. It does care about elegance, which makes sense if you think about the two or three facts that we have about the conditions for the emergence of the language faculty. All this hangs together pretty reasonably.

6) *What is complexity in language? And in thought?*

To answer the question about either language or thought, we first have to say what they are. You can't answer questions about something when you have no idea of what it is. In the case of language, we can say some things, there are some ideas about what the essential nature of languages is. What is complexity in language? If we adopt the Galilean guideline for the sciences, complexity in language is whatever we don't understand: if you don't understand something, it is complex. What you are trying to do is to show that it is your failure to understand that gives the impression of complexity: that is the nature of rational inquiry, science in particular. So there is plenty of complexity in language; pick anything random out of a page and it is very complex, which simply means that we don't understand enough. But that is true of everything you look at in the world. That is why physicists don't take video tapes of what is happening in the worlds and try to develop theories: this is way too complex. In fact, there is a standard joke in mathematics, that the only numbers are 1, 2, 3 and infinity, because anything else is just too complicated to study. In the case of physics, hydrogen you can study, helium maybe, but, if it gets bigger than that, you give it to the chemist and they will worry about it. If an organic molecule is too big or complicated for a chemist, you give it to the biologist; if it is too big for the biologist, you give it to historians; so it goes. You can study simple things, and when they get too complicated it gets harder, which means that you don't understand enough. That is the nature of things.

That is language. What about thought? You can't do this for thought, until someone tells you what thought is. What is thought? We can say a couple of things about those aspects of thoughts that are expressed in language, but then we are talking about language. What about those aspects of thought that aren't expressed in language? What do we know about them? Actually what we know about them is mostly from introspection. I am sure that I am not the only person in the world who has the experience of knowing what I want to say but not thinking of how I can say it: "I know what I mean but I can't find the way to say it", and you try something that didn't work, maybe you try something else, maybe it is a little better, maybe you end up with some complicated paraphrase because you couldn't figure out the right way to say it. That is a normal experience. Experience is like that. It tells you something: that there are a lot of things going on beyond the level of consciousness that we try to move to consciousness and even to the external world, often failing, which means that there is a lot of thought going on, and we have no grasp of it. Until there is some way to capture these things that are beyond the level of consciousness, which is probably almost everything in our cognitive world, there is nothing to say. There is other evidence about this that has been misinterpreted, I think. In the last couple of years ago, there were experiments showing that, when people make decisions, for example when I decide to pick up this cup, milliseconds before I make the decision, there is activity in the brain in the areas where you are going to act, i.e., milliseconds before I make the decision, the motor areas of the brain are already organized to pick the cup up. That evidence was used widely to conclude that this shows that we don't have free will¹⁰. But this doesn't show anything of this sort. This just shows that decisions are unconscious. We all know that, if we think for a minute: of course decisions are unconscious. Some of them reach the level of consciousness, some of them we can't even act on, but there is a lot there going on unconsciously, probably everything of interest, and we don't know how to deal with it. In principle it could be studied, maybe some day brain scientists reach the point where they can say something about this: it doesn't seem beyond the possibility of inquiry. But we have to overcome some dogmas. There is one dogma that has hampered psychology back hundreds of years: that is the dogma that contents of thought are accessible to consciousness. It is hard to find anyone in the history of psychology or philosophy who has doubted this, even Freud. Freud talked

¹⁰ To this respect, see the studies of the neurophysiologist Benjamin Libet on voluntary movement (Libet 1985). On the implications of this type of evidence for the notion of free will, see Pietrini & Bambini (2009).

a lot about the unconscious, but the whole Freudian system is based on the assumption that you can tease out what is unconscious. That is the point of psychoanalysis: they try to make what is unconscious, conscious. But if you try to find somebody who questioned that, it is not easy; the only person I have ever found is Vico. This dogma is almost certainly false. We know that from our own experience. Going back to the question, until we have something to say about thought, we are not going to be able to ask questions about its complexity, its simplicity, its properties, or anything else.

7) *What comes first? Language or thought? Do they share a recursive structure?*

It is the same question: until we have something to say about thought, we are not going to be able to give any serious answer. The only answer we can give is the one I have already mentioned: you simply pay attention to your own internal cognitive activity, speaking, planning, deciding, and so on. And I think you invariably come to the conclusion that something is going on beyond the level of consciousness. Probably, everything interesting is going on there. That is thought, if you like, and, whatever it is, it is coming before language, before deciding, before acting: it is just going on inside. But what is it? Well, it is a task for the future.

8) *What is the place of context in language functioning? For example, metaphors are context dependent. Do they lay outside the language faculty?*

Interpretation of metaphors, like interpretation of everything else, is context-dependent and lies outside the language faculty, specifically, unless by the language faculty you mean all of our cognitive abilities¹¹. So, yes that would be true of metaphors, but it is also true of non-metaphors. Whatever you interpret in literal sentences goes way beyond the language faculty. Again, there are a lot of dogmas about this.

9) *Why is it so difficult to explain that the difference between human and animal language lies the syntax?*

That relates to the previous question. First of all, it is difficult to explain anything about language, because people have religious attitudes towards their language, religious dogmas about their language, as about almost anything else that is deeply personal. So you just feel you know everything: “you can’t tell me anything, I can speak perfectly and I don’t have any problem, there can’t be any difficulties about it, so stop bothering me with all your complicated theories”. That is very normal. You take a look at modern philosophy of language and it is full of this. People spell out their untutored intuitions about language and think they are saying something, because how could it be complicated? You could say the same about vision. What could be complicated about vision? “I see things, I don’t have any problems with it”. If you try to explain to people that the way you see things is because of quite complex computations that take place in the visual system, there is interesting experimental work which with you can sometime convince people. Somehow this goes back to unconscious knowledge. The mammalian visual system, humans and other mammals that have been studied, has a way of interpreting objects in motions as rigid¹². In standard experiment, if you give a person tachistoscopic presentation, like a screen with a couple of dots on it, maybe three dots, if you give them three or four successive presentations of three dots on the screen, what you see is a rigid object in motion. You can’t help that: that is just the way the visual system works, which is kind of strange in a way. In the whole history of mammals, going as far back as you like, there was never any experience with rigid objects: rigid objects come in what is called “carpentered universes”, in modern universes where people construct things. If you are walking around in a forest, there aren’t any rigid objects. But our eyes and visual systems in all other mammals are constructed so that all of this completely unconscious activity is going on inaccessible to consciousness, which is giving a kind

¹¹ On the constellation of cognitive abilities involved in interpreting metaphor and context-dependent meanings in general, see Bambini *et al.* (2011).

¹² See the work on visual perception by David Marr (1982) and Shimon Ullman (1979), and specifically Ullman’s rigidity principle.

of a framework for visual perception. Try to explain that to somebody about cognitive abilities like language and they would just resist it: "It can't be. I know everything, don't bother me". It is hard to explain anything about language, unless somebody is willing to take on the attitude of the standard scientist: my intuitive judgments don't mean anything, maybe my intuitive judgment is that a heavy rock falls faster than a small rock; if I want to be serious, I put aside this intuition because it has been disproved. Unless you can make that leap, nothing can be explained.

What about the special role of syntax? The fact that matters is that human syntax has no counterpart in the animal world, none. No one has ever found any remotely like it. This fundamental combinatorial process that I mentioned simply appears nowhere else, at least anybody has been able to detect. But is that the difference between animals and humans? Well it is a difference, it is only one of many. If you take even the simplest word of language, the simplest one you like, "river", "cup", "person", whatever it is, there is absolutely no analogue in animal communication. Animal communication has essentially a 1 to 1 association between a symbol and some physically detectable set of circumstances. Take monkey calls: monkey might have five calls. One of them is reflexively produced when leaves are moving: it is a warning call, we interpret that as meaning "predator is coming". Leaves are moving, this call comes out: it is like seeing a rigid object in motion. Another call comes out because of some hormonal activity: we interpret its meaning as "I am hungry". In every animal communication system that is known the symbols are like that. Human words are nothing like that, not remotely even. You cannot give the physical characterization of what it means to be a river, or a person, or a cup. All of these concepts, even the simplest ones, depend on internal interpretations in terms of function, design, intentions, psychic continuity, all sorts of things. And every infant grasps this instantaneously, and in fact fairy tales are based on it. Standard fairy tale has some prince turned by the wicked witch into a frog and he stays this way until the beautiful princess kisses the frog that becomes the handsome prince. During the interim period the physical characterization of the object was a frog, with all physical properties of the frog. But the child understands it as the prince: it just happens to look like a frog. That means that the concept of person is based on an intuitive notion of psychic continuity, which has no physical characterization, and that is instantaneous: no child has ever been confused about this. That is true for every word you look at. So, it is true that the syntax is different for the humans and animals, but so it is for everything else about language, including the meaning of the simplest words. This is another dogma, very resistant to reason and evidence, which has prevented the understanding of this. It is sometimes called the "referential doctrine": the idea that there is a relation between words and things, an association between a word and the thing, that runs through philosophy of language and history of inquiry into the subject in recent years. That is just flatly false. There is no such relationship. In fact that was known to Aristotle, but it has been forgotten. If you look back at Aristotle, he asks: what is a house? In the Aristotelian framework, a house consists of the interaction of two substances: the one is "matter" – house is something that is built out of bricks and wood and so on, its materials aspect – and the other is what is called "form", design and function – it is used for a place for people to live in. If form and matter coincide, you get a house. But form is not physically detectable. And in fact the Aristotelian definition, though is correct as far as it goes, it just doesn't begin to reach the meaning of house. But that is the right idea. That was understood pretty well until the 17th century, like a lot of things that were understood then and that are almost totally forgotten. Until those insights are recovered, it is going to be impossible to study this question.

But to go back to the question itself, it is true that syntax sharply separates human language from animal communication, so does the interpretation of the simplest atomic elements, so does everything else. Animal communication systems are – as the term indicates – used for communication. For human language that just doesn't seem to be a central property. It is certainly one of things you can do with language, just as you can watch television with your eyes. But that

is not a core property and there is plenty of evidence for that, of the kind I mentioned. So, in every respect it is just some novel thing in the organic world. I mentioned this morning Descartes: I think he was basically right in thinking that this is the dividing between humans and the rest of the organic world, something that happened very recently¹³.

10) What is the difference between syntax in music and syntax in language? If they are alike, is it plausible that they have coevolved?

It is very plausible. Language is universal among humans. There is no human group that has been found that doesn't have a language very much like ours. In fact, languages are even invented by children with zero evidence. There are some remarkable cases. I will mention one case which is well studied. It is now known – it wasn't known 50 years ago – that sign languages are almost identical to spoken languages: they have the same structures, same rules, same acquisition rate, same neural representation even, which is a little surprising because they are visual. There is one well studied case of a group of three cousins, in Philadelphia, whose parents were indoctrinated into the prevailing so-called oralist tradition: deaf children should not be permitted to use gesture, they have to learn lip reading. The idea behind this is crazy but it was the prevailing ideology in raising children with deafness. These parents were so deeply indoctrinated that they never made gestures, they walked around the house with their hands behind their backs so that the kids wouldn't see any gesture. These three kids played together and it turned out, around the age of 3 or 4, that they invented a sign language: they developed a sign language and they were just using it. They were immediately taught American Sign Language, but the language was studied and it was about the same of any other language at that developmental level¹⁴.

Actually, the first study of this was one of a friend of mine, back when I was a graduate student, Eric Lenneberg, who went on to develop the field of biology of language¹⁵. When he was a graduate student, back in the early 1950s, he was getting interested in the acquisition of language under a variety of conditions, and particularly language of the deaf. He went to visit a prestigious school for the blind and deaf in Boston, called the Perkins School¹⁶. It was oralist, like everything was: no sign language, just lip reading. He just sat in on a class and he noticed that, as soon as the teacher turned to the blackboard, the kids started signing to each other. Obviously the kids had invented a sign language, somehow. There was no way to study it at that time, but it is very likely, it is spontaneous, you can't help it.

Music is the same. As far as the anthropological evidence indicates, no human group has been found that hasn't developed a pretty complex system of music. It may not be tonal music. It might be rhythm based or something else. But some kind of system of music seems to be universal, which is very curious because what function does it have? Why should there be music? A similar question arises about arithmetical knowledge. That is also universal. This is something that bothered Darwin and Wallace, founders of evolutionary theory, because it seems to violate the principle of natural selection. How could arithmetical knowledge have evolved, since it has never been used? It has only been used in a tiny period of human history, and only among very few people. How come that everybody has it? And where is the good evidence that this is universal? That looked problematic. The only sensible answer that has been proposed is that all of these are offshoots of the same cognitive system. Some cognitive system that emerged, which led to arithmetic, music, language, and any

13 In the morning of September 15, Chomsky gave the inaugural lecture at the Institute for Advanced Study, IUSS, on *Language and limits of understanding*, considering also Descartes's view on language and human creativity.

14 On deaf children spontaneously developing a sign system, see Goldin-Meadow & Feldman (1977) and Goldin-Meadow (2003).

15 On the biological foundation of language development, including the condition of congenital deafness, see Lenneberg (1967; 1969).

16 The Perkins School for the Blind was established in 1829. Famous students at the Perkins School were Anne Sullivan and Helen Keller.

other behavioral pattern that has this basic properties and is universal. There are some reasons to believe that the combinatorial operation that I mentioned before actually does underlie all of them. For arithmetic, it is pretty easy to show, for music it is more complicated, but there is some work on it. In the last 40 years there has been a certain amount of work on the syntax of music, trying to relate it to the syntax of language. The first major studies were done by Leonard Bernstein, American well known composer and conductor. There was a lecture in Harvard, back in the early 1970s¹⁷. Since then there has been a certain amount of work. One of my colleagues in my department, David Pesetsky, is one of the people that has done the main work on this¹⁸. It is an interesting topic. You can see how it ought to end up. But trying to show that what ought to be true is true, is never an easy task.

11) What is your position about the new theories of embodied language? In particular, what do you think about the relationships between the motor system and semantics?

That is a quick one. I don't have any position. "Semantics" is a pretty loose term. If by "semantics" you mean everything involved in interpretation of language, then it would be pretty strange if it didn't have some relation to motor systems and everything else. What that relation might be, I am not aware of any more than superficial observations relating them. I am afraid I have to say the same thing about embodied language: that is another popular topic. If you think about things like emotions and reactions to things and so on, what has that to do with language? A lot, it has a lot to do with the use of language, trivially, but does it have to do with the structure of language, the principles that determine the infinite structure of an interpretative expression? As far as I am aware, nobody has thought of a connection and it would be pretty hard to imagine one. They just seem quite different systems¹⁹.

17 Bernstein delivered six lectures at Harvard in 1973, broadcasted in 1976 and available also as a book (Bernstein 1976).

18 As a representative reference, see Katz & Pesetsky (2009), which follows the path inaugurated by the seminal work of Ray Jackendoff and Fred Lerdhal (Lerdhal & Jackendoff 1983).

19 A discussion on embodied cognition theories with respect to the language faculty can be found in Tettamanti & Moro (2012).

REFERENCES

- Bambini, V., Gentili, C., Ricciardi, E., Bertinetto, P.M., and Pietrini, P. (2011), "Decomposing Metaphor Processing at the Cognitive and Neural Level Through Functional Magnetic Resonance Imaging", *Brain Research Bulletin* 86, 2001, pp. 203-216;
- Bernstein, L. (1976), *The Unanswered Question: Six talks at Harvard*. Cambridge, Mass., Harvard University Press, 1976;
- Chesi, C. (2012), *Competence and Computation: Toward a Processing Friendly Minimalist Grammar*, Padova, Unipress, 2012;
- Chomsky, N. (1988), *Language and Problems of Knowledge: The Managua Lectures*. Cambridge, Mass., MIT Press, 1988;
- Chomsky, N. (1995), *The Minimalist Program*. Cambridge, Mass., MIT Press, 1995;
- Chomsky, N. (2004), *The Generative Enterprise Revisited: Discussions with Riny Huybregts, Henk van Riemsdijk, Naoki Fukui, and Mihoko Zushi, with a new foreword by Noam Chomsky*, Berlin, Mouton de Gruyter, 2004;
- Chomsky, N. (2011), "Problems of Projection", to appear in *Lingua*;
- De Palma, A. (1974), *Linguaggio e Sistemi Formali*, Torino, Einaudi, 1974;
- Miller, G.A. (1956), "The Magical Number Seven, Plus or Minus Two: Some Limits on Our Capacity for Processing Information", *Psychological Review* 63, 1956, pp. 81-97,
- Moro, A. (1996), "Virtual Conceptual Necessity: la semplificazione della grammatica generativa nei primi anni novanta", *Lingua e Stile* 30, 1996, pp. 637-674;
- Moro, A. (2008), *Boundaries of Babel. The Brain and the Enigma of Impossible Languages*, Cambridge, Mass., MIT Press, 2008;
- Goldin-Meadow, S. (2003), *The Resilience of Language*, New York, Psychology Press, 2003;
- Goldin-Meadow, S. and Feldman, H. (1977), "The Development of Language-Like Communication Without a Language Model", *Science* 197, 1977, pp. 401-403;
- Goodman, N. (1954), *Fact, Fiction and Forecast*, Cambridge, Mass., Harvard University Press, 1954;
- Katz, J. and Pesetsky, D. (2009), "The Identity Thesis for Language and Music", <http://ling.auf.net/lingBuzz/000959>;
- Lenneberg, E.H. (1967), *Biological Foundations of Language*, New York, Wiley, 1967;
- Lenneberg, E.H. (1969), "On Explaining Language", *Science* 164, 1969, pp. 635-643;
- Lerdahl, F. and Jackendoff, R. (1983), *A Generative Theory of Tonal Music*. Cambridge, Mass., MIT Press, 1983;
- Libet, B. (1985), "Unconscious Cerebral Initiative and the Role of Conscious Will in Voluntary Action", *Behavioural and Brain Sciences* 8, 1985, pp. 529-566;
- Marr, D. (1982), *Vision*, San Francisco, Freeman, 1982;
- Pietrini, P. and Bambini, V. (2009), "Homo Ferox: The Contribution of Functional Brain Studies to Understanding the Neural Bases of Aggressive and Criminal Behavior", *International Journal of Law and Psychiatry* 32, 2009, pp. 259-65;
- Smith, N.V. and Ianthi-M.T. (1995), *The Mind of a Savant: Language learning and modularity*, London, Blackwell, 1995;
- Tattersall, I. (1998), *The Origin of the Human Capacity*, New York, American Museum of Natural History, 1998;
- Tettamanti, M. and Moro A. (2012), "Can Syntax Appear in a Mirror (system)?", *Cortex* 48, 2012, pp. 923-935;
- Ullman, S. (1979), *The Interpretation of Visual Motion*. Cambridge, Mass., MIT Press, 1979.

