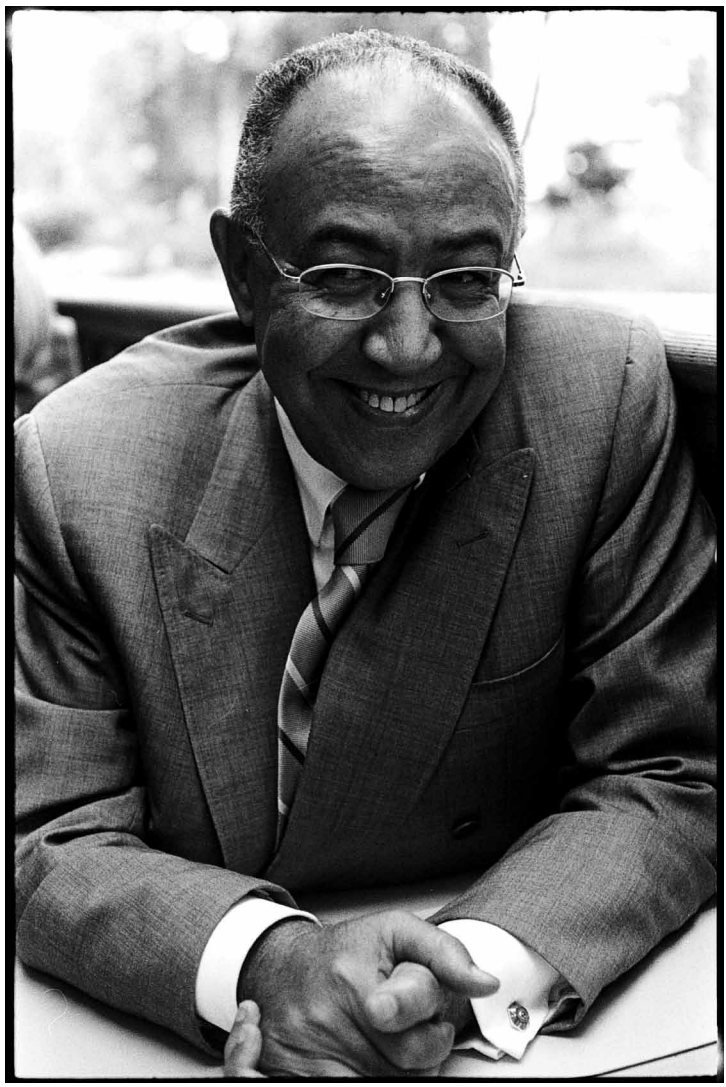
THEORIES AND SOCIAL MECHANISMS



Mohamed Cherkaoui, Paris, 2007
Photograph by Peter Hamilton

THEORIES AND SOCIAL MECHANISMS

Essays in Honour of Mohamed Cherkaoui

VOLUME I

Edited by
Gianluca Manzo

THE BARDWELL PRESS

Introduction and editorial arrangement © Gianluca Manzo, 2015
Individual contributions © the authors, 2015

First published 2015

All rights reserved. Apart from any fair dealing for the purposes of research or private study, criticism or review, as permitted under the UK Copyright Designs and Patents Act, 1988, no part of this publication may be reproduced, stored or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording or otherwise, without the prior permission in writing of the publishers.

Published by:

The Bardwell Press
Tithe Barn House
11 High Street, Cumnor
Oxford, OX2 9PE
United Kingdom

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

ISBN: 978-1-905622-46-7 (Set of 2 volumes)

Typeset by The Bardwell Press, Oxford, UK
Printed in the United Kingdom by Berforts Information Press

CONTENTS

VOLUME I

List of Contributors xi

INTRODUCTION

Why Read Cherkaoui's Work? A Road Map
Gianluca Manzo 1

PART ONE
History of Social Thought

1. Sociodicy: Some Comments on Mohamed Cherkaoui's Durkheimian Approach to Social Emergence
Salvador Giner 41
2. Back to the Origins of the Sociology of Generative Mechanisms
Dominique Guillo 47
3. Engineering the Welfare State: Gunnar Myrdal, Social Science and the Politics of Unintended Consequences
E. Stina Lyon 61
4. "Orientation to Others": A Central but Forgotten Concept in Max Weber's Sociology
Richard Swedberg 91

PART TWO
Epistemology of the Social Sciences

5. Generative Mechanisms and Decreasing Abstraction
Jack Birner 115
6. Methodological Individualism and the Notion of Levels of Social Life
Pierre Demeulenaere 141
7. Macro Mechanisms: Is Sociological Explanation Possible in the Absence of Action Theory?
Loris Di Giammaria 163
8. The Strategy of Mechanism-Based Research
Diego Ríos 185
9. The Plurality of Explanatory Games
C. Mantzavinos 201
10. Reduction That Really Matters: Can We 'Reduce' from Macro to Micro in the Social Sciences (And Should We)?
José A. Noguera 213
11. Aggregation Paradoxes and Non-Commutativity
Dominique Raynaud 229

PART THREE
Social Theory

12. Rituals as Mechanisms
Filippo Barbera & Nicola Negri 251
13. The Four Rules of "Good Sociology"
Raymond Boudon 277
14. Purposive Actions and Their Unintended Consequences
Jacques Coenen-Huther 303
15. Instrumental, Axiological Rationality and the Explanation of Norms: Cherkaoui's (and Boudon's) Critique of Rational Choice Theory and Its Ability to Explain Norm Emergence
Karl-Dieter Opp 319

16. Social Interdependence, Rights and Normative Dynamics
Emmanuel Picavet 343
17. Functionalism, Field Theories, and Unintended Consequences
Stephen P. Turner 365

VOLUME II

PART FOUR
Education and Social Mobility

18. Mohamed Cherkaoui: His Thoughts and His Impact and Validity in Argentina
Miriam Aparicio 3
19. Changes in the French Education System and Inequality of Opportunity Issues: Going Beyond Preconceptions
Nathalie Bulle 21
20. School Dropout and Related Assessments: The Case of Morocco
Brahim Chedati 41
21. Equity, Equality and Microjustice
Michel Forsé 61
22. Social Mobility, Precariousness and Politics in 2012
Nonna Mayer 81

PART FIVE
Status Distributions

23. Invisible Community and Peerless Science: Case Studies from Morocco and India
Michel Dubois 101
24. Scientists and Fields: A Theoretical Case Study
Thomas J. Fararo & John Skvoretz 155

25. Social Recognition and Success in the Contemporary Visual Arts: A Sociological Perspective on Rankings of the "Top 100 Artists in the World" and the Uneven Distribution of Success Among Nations
Alain Quemin 181

PART SIX
States and Markets

26. International Relations Analysis With Game Theory
Pierre Allan & Cédric Dupont 209
27. Norms and Derogation in Morocco's Public Finances: A Durkheimian Explanation
Noureddine Bensouda 223
28. Equity and Self-determination in Western Sahara: An Alternative Path to an Enduring Compromise
Rachid el Houdaïgui 237
29. Global Domination of Finance Beyond Pioneers' Expectations
Jacques Lautman 269

PART SEVEN
Social Movements

30. Terrorism and Rationality
Gerald Bronner 289
31. Local Truth Mechanisms: About Moroccan Political Consensus
Jean-Noël Ferrié & Baudouin Dupret 311
32. Social Protests in Morocco: Riot, Emotions and Reasons
Abderrahmane Rachik 327
33. On the Authority of Disciples: Essay on the Rise of Islamic Fundamentalism and Its Transformation
Hassan Rachik 349
34. Invisible Codes of Terror: A Sociology of Political Violence
Domenico Tosini 367

35. Irony and Paradox in Sociological Explanations: Mohamed Cherkaoui on Revolutions and Good Intentions
Bryan S. Turner 387

CODA

Cross-cutting Fields and Social Circles: A Conversation with Mohamed Cherkaoui
Mohamed Cherkaoui
Interviewed by Gianluca Manzo & Peter Hamilton 407

LIST OF CONTRIBUTORS

Pierre Allan
University of Geneva.

Miriam Aparicio
National Council for Scientific Research (CONICET), Argentina.

Filippo Barbera
University of Turin and Collegio Carlo Alberto—Moncalieri.

Noureddine Bensouda
Treasurer General of the Kingdom of Morocco.

Jack Birner
University of Trento; University College Maastricht. Konrad Lorenz Institute for Evolution and Cognition Research.

Gerald Bronner
University of Paris-Diderot—LIED (UMR CNRS).

Raymond Boudon
Académie des Sciences Morales et Politiques, Paris.

Nathalie Bulle
CNRS, Paris; GEMASS, Paris.

Brahim Chedati
Centre d'Orientation et de Planification de l'Education (COPE), Rabat, Morocco.

Jacques Coenen-Huther
University of Geneva.

Pierre Demeulenaere
University of Paris-Sorbonne.

Michel Dubois
CNRS, GEMASS, Paris.

Cédric Dupont
Graduate Institute of International and Development Studies, Geneva.

Baudouin Dupret
CNRS, CJB (Rabat, Morocco).

Thomas J. Fararo
Distinguished Service Professor Emeritus of Sociology, University of Pittsburgh.

Jean-Noël Ferrié
CNRS, CJB (Rabat, Morocco).

Michel Forsé
CNRS, Paris.

Loris Di Giammaria
Sapienza University of Rome.

Salvador Giner
University of Barcelona.

Dominique Guillo
CNRS (GEMASS UMR 8598, CNRS-Paris IV).

Rachid el Houdaïgui
Faculty of Law, Abdelmalek Essaadi University, Tangiers.

Jacques Lautman
Université Aix-Marseille.

C. Mantzavinos
University of Athens.

Gianluca Manzo
GEMASS—CNRS & Paris-Sorbonne.

Nonna Mayer
Centre d'Études Européennes de Sciences Po—CNRS.

Nicola Negri
University of Turin.

José A. Noguera
Grupo de Sociología Analítica y Diseño Institucional (GSADI), Department of Sociology, Universitat Autònoma de Barcelona.

Karl-Dieter Opp
University of Leipzig / University of Washington (Seattle).

Emmanuel Picavet
Université Paris-1 Panthéon-Sorbonne.

Alain Quemin
Université Paris-8 / Institut d'Etudes Européennes.

Abderrahmane Rachik
King Abdul-Aziz Al-Saoud Foundation, Casablanca, Morocco.

Hassan Rachik
Hassan II University, Casablanca.

Diego Ríos
Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Buenos Aires, Argentina. UNTREF, Buenos Aires, Argentina.

Dominique Raynaud
PLC, Université Pierre-Mendès-France (Grenoble).
GEMASS CNRS UMR 8598 (Paris).

John Skvoretz
Professor of Sociology, University of South Florida.

E. Stina Lyon
London South Bank University.

Richard Swedberg
Cornell University.

Domenico Tosini
Department of Sociology and Social Research, University of Trento, Italy.

Bryan S. Turner
The Graduate Center, The City University of New York

Stephen P. Turner
Department of Philosophy, University of South Florida.

WHY READ CHERKAOUI'S WORK?

A Road Map

Gianluca Manzo
GEMASS—CNRS & Paris-Sorbonne

INTRODUCTION

Through a series of fourteen well-articulated definitions, in his famous essay "Why read the classics?", the Italian writer and literary critic Italo Calvino (1991 [1999]) identified several distinctive features of a classic, among which are: 1) the fact of being "re-read" more than "read" (*ibid.*: 3); 2) its capacity to set conceptual, value, and aesthetic standards (*ibid.*: 4); 3) its being open to continuous re-interpretation (*ibid.*: 5–6); 4) its capacity to look "original, unexpected, and innovative" when one reads it compared to "hearsay" (*ibid.*: 6); 5) its capacity to raise criticism and opposition (*ibid.*: 7). The essays collected in the present two-volume *Festschrift* in honour of Mohamed Cherkaoui suggest that his books and articles share many of these features.

In his provocative essay, Calvino also defended the argument that it is not the "antiquity, style, or authority" of an author that make it a classic but his or her capacity to be part of a "cultural continuum". That capacity, he notes, should be measured through the ease with which "its place in the genealogy of classic works" (*ibid.*: 7) can be recognized.

In this respect, too, Cherkaoui's scientific production can easily be classified. It is clearly part of the middle-range-theory tradition in

sociology (see Merton, 1967; 1987). Similarly to James Coleman, Raymond Boudon, Thomas Fararo, John Goldthorpe and Peter Hedström—to mention just a few authors with whom Cherkaoui often dialogues in his writings—, Cherkaoui's oeuvre is animated by a deep explanatory ambition and combines quantitatively sophisticated empirical research, theoretical analyses, epistemological reflection and the study of of classics. A distinctive feature of Cherkaoui's scientific output is the diversity of substantive topics, methods, and investigations in the history of sociology this work covers. The variety of national intellectual communities and disciplinary backgrounds from which the contributors to the present *liber amicorum* come well reflects this variety.

With the aim of helping readers to navigate Cherkaoui's works and, at the same time, to perceive the coherence behind the apparent heterogeneity of the thirty-five contributions that follow, in this introductory essay, I survey the main research areas in which Mohamed Cherkaoui was involved during his intellectual career. My aim is not exhaustiveness. Cherkaoui's bibliography is so vast that it would be impossible to do full justice to it. More modestly, I will seek to identify, and, to some extent, to periodize, Cherkaoui's major thematic focuses and lines of development. The focus is on what I regard as the major contribution of each (set of) article(s) and book(s), not on the integrality of details and results they contain. Incidentally, I will also try to show how Cherkaoui's contribution fits, and contributes to, past and more recent debates. It is a road map, or, if you prefer, a set of guidelines, that I would like to draw. The limitations of this operation are discussed in the concluding section. Let me also note that, on purpose, I leave aside contextual and biographical details. The previously unpublished interview Mohamed Cherkaoui graciously gave us (see Cherkaoui, 2015, present volume) thus is a necessary (and extremely pleasant to read) complement to the present introduction.[1]

QUANTITATIVE METHODS

In the sixties and seventies, the application of statistical methods for the analysis of empirical data in sociology was still in its infancy (for an overview, see Raftery, 2001). Mathematical sociology, too, as a specific

field of research, was only emerging (for an overview, see Edling, 2002). Mohamed Cherkaoui belonged to this minority of scholars who believed in the importance of statistics and mathematics to describe and explain social phenomena.

Let me mention first his analysis of the dependence of students' achievement in mathematics on class size and teaching length as a function of several other variables like track, social class or sex (see Lindsey and Cherkaoui, 1975). To describe this association (across six different countries) Cherkaoui estimated a series of linear multiple regressions specifying different types of interaction effects among the controlling variables and assessed when these effects were statistically significant (*ibid*.: appendix). It should be noted that this analysis was performed after correcting for the asymmetric nature of the dependant variable, through a Box-Cox transformation, a topic that Cherkaoui also developed separately (see Cherkaoui, 1976). When one notes that even today marginal effects are often given priority over interaction effects (for this critique against regression-like analysis, see, for instance, Sorensen, 1998), and that variable transformations are still common only among a minority of advanced users (see Gelman and Hill, 2009, chapter 4), the sophistication of Cherkaoui's analysis becomes apparent.

A similar remark holds for Cherkaoui's critical analysis of the literature on how to measure social classes (Cherkaoui and Lindsey, 1977), without a doubt one of the longest-lasting debates in the quantitative sociology of social stratification (for a recent overview, see, for instance, Hauser and Warren, 1997; Chan and Goldthorpe, 2007). In this respect, Cherkaoui provided us with a deep critical analysis of the methodological foundations of the variety of socio-economic indexes and prestige scales available in the seventies, which he attacked for relying on a continuous view of social structure and statistical tools based on the hypothesis of linearity. On this basis, he argued in favour of a theoretical approach that takes into account the existence of discontinuities among social groups and suggests statistical procedures that would describe the observed complexity of the social achievements of these groups. In particular, on the one hand, he suggested an ingenious form of variance analysis for data violating the assumption of normality and that pays attention to the interaction effects among several independent variables

(see Cherkaoui and Lindsey, 1977: 256–261); on the other hand, he proposed a way to analyse case over-representation in contingency tables in line with the nascent statistics for categorical data (*ibid*.: 261–264).

The term and the explicit reference to the formalism of "log-linear analysis" overtly appeared in Cherkaoui's (1979a: 98–102, 104) first comprehensive treatment of educational achievement in comparative perspective (see next section), in particular where he studied the cross-country variation of the complex relation between performances in mathematics, types of schools, tracks, and social classes. Even more refined use of log-linear models is found in Cherkaoui's (1982: 130–132) major contribution to sociology of education (see next section), in particular where he analyses the complex interactions between being a candidate to the *baccalauréat*, and, on the other hand, his or her age and social class. In this application, Cherkaoui also shows how to compute odds ratios from the estimated parameters, thus making the results more interpretable and sociologically meaningful. By stressing the importance of log-linear models for the analysis of categorical data, a research area in social statistics that would later become central for the methodology of contemporary sociology of social stratification (see, for an overview, Goodman, 2007)—Cherkaoui incontestably was ahead of his time.

This is testified by at least two other methodological analyses that appear in Cherkaoui's (1979a, 1982) book-length studies of educational systems. First, in his regression-based analyses of the extent to which the length of teaching, or the quality of teachers, impacts on educational achievement, or on other dependent variables like "aspirations", Cherkaoui not only pays systematic attention to the interaction effects among these variables but also attempts to describe the variations in the functional forms relating variables across population sub-groups (see, Cherkaoui, 1979a: chapters 4, 5, and 6, respectively). As suggested by the "contour plots" Cherkaoui drew to express the results of these complex models (see in particular, figs. 5.1.1–5.2.6, 6.1.2, and 6.2.2), this analysis followed the principles of the emerging "response surface methodology", which today is a flourishing approach widely applied in chemical and biological research as well as in the industrial sector (see, for an overview, Khuri and Mukhopadhyay, 2010).

Second, to describe the increasing participation in upper secondary school in France between 1950 and 1978, Cherkaoui (1982: chapter 1) took inspiration from time series analysis in economics and assessed the fit of several functions (including Gompertz's function) to describe the observed trends, a procedure that is nowadays common among practitioners of "growth curve modelling" (see Panik, 2014). Cherkaoui did not stop at description, however. In his analysis of the temporal changes in the age distribution of *baccalauréat* candidates, for instance, Cherkaoui employed what we now call "stochastic processes theory" (see, among others, Gallager, 2013), and tested a variety of probability distributions (included quite refined ones such as Poisson distributions with contagious process) against actual data, the rationale behind this procedure being that the formal properties of the probabilistic model with the best fit may reveal the nature of the underlying social mechanisms at work (see Cherkaoui, 1982: chapter 4).

Thus, Cherkaoui's early empirical studies in the field of sociology of education relied on some of the most refined quantitative procedures available in the seventies. This should not lead readers to ignore the fact that a systematic interest in a variety of quantitative and formal tools persists throughout Cherkaoui's subsequent scientific research. These include, for example, the mathematical modelling based on differential equations that Cherkaoui employed to formalize Durkheim's theory of social change and integration (see, respectively, Cherkaoui, 1981a [2005: 45–49, 163–167; 2008: 46–64]). He has also constantly paid attention to numerical simulation, which he regarded as a powerful support for theoretical elaboration and useful complement to statistical analysis (see Cherkaoui, 2001; 2006a [2007a: 3–4]). Let me finally mention Cherkaoui's long-standing interest in game theory, which he most recently applied to frame the complex relationship between Morocco and Algeria with respect to the issue of the independence of Western Sahara (see Cherkaoui, 2007b: chapter 1: pp. 10–13 and footnote 3).

SOCIOLOGY OF EDUCATION

As we have seen, Cherkaoui's early research in the field of sociology of education remains remarkable, even today, for its methodological

refinement. To understand the substantive originality of this work, it is important briefly to summarize the debates to which it aimed to contribute.

One of the deepest social changes affecting western societies in the sixties concerned the increasing investment in secondary education families started to make. The question of whether these behaviours was able to generate better social opportunities for all social groups thus became one of the most urgent empirical questions in sociology. Several studies, based on different theoretical and methodological perspectives, converged to a negative reply denying that the educational system *itself* can play an *autonomous* role in reducing social inequalities. In the US, Coleman *et al.*'s *Equality of Educational Opportunity* (1966) focused on the distribution of resources within and across American schools and found that school variables (like types of educational activities, libraries, social composition, teachers' years of experience, etc.) played a very minor role, compared to pupils' attributes (like family background), in determining educational achievements. Jencks *et al.* (1972) observed that education contributed only modestly (statistically) to explain individuals' opportunities on the job market, thus concluding that job search is driven more by chance than by formal education. In Europe, Bourdieu and Passeron (1970) argued that school can only reproduce social inequality because successfully going through the educational system requires unequally distributed social and cultural skills that the school in fact does not provide. Boudon (1974) stressed that, since the occupational system tends to evolve more slowly than the educational one, more educated individuals will not necessarily experience upward intergenerational social mobility.

Cherkaoui's (1979, 1982) book-length analyses of modern educational systems established a systematic dialogue with this literature with the goal of demonstrating that school organization in fact plays a specific role in generating differential educational outcomes across social groups, and that, to understand how educational systems evolve, we need to acknowledge and theorize their *relative* autonomy.

More particularly, Cherkaoui's (1979a) first systematic contribution to the sociology of education[2] can be read as a re-examination of Coleman's main conclusions concerning the low impact of school

characteristics on educational stratification. With the aim of demonstrating that Coleman's results do not necessarily hold everywhere, Cherkaoui used data from the *International Project for the Evaluation of Educational Achievement for pupils* (aged around 13) and compared seven countries (namely, France, Belgium, England, Scotland, Sweden, US, and West Germany). Among school variables, he essentially focused on tracking (general *versus* technical) and type of procedures for pupil selection. Relying on multiple variance analysis as well as techniques for categorical data (see previous section), Cherkaoui established that, while pupils' social class has a strong impact on educational achievement (in mathematics) in all countries, this effect significantly interacts with school characteristics varying across countries. In particular, Cherkaoui found that the stronger the selectivity and more visible track segmentation, the weaker was the effect of social class. This result, Cherkaoui argues, suggests that Coleman's results concerning the low impact of school variables on educational inequalities do not necessarily hold in European educational systems, in particular when strong selection and internal differentiation operates at the earlier educational stages (see, in particular, *ibid*.: chapter 3). As recent literature shows (see, for instance, Hanushek and Woessmann, 2006), forty year later, the issue raised by Cherkaoui continues to stimulate debate, and, despite methodological advances, a definitive answer has still to be found.

Cherkaoui's (1982) second major contribution to the sociology of education developed further the idea of the relative autonomy of school organization in shaping educational inequalities. In particular, Cherkaoui quantitatively described several changes in the French educational system after World War II and, for each of them, formulated micro-based theoretical explanations.[3] Aggregate consequences that can logically be derived from these explanations are systematically confronted with consequences derivable from alternative accounts focusing on macroscopic (economic or institutional) changes that are exogenous to the education system. Cherkaoui demonstrated that the statistical structure of data describing the relevant changes is incompatible with the latter group of explanations. Ultimately these results are used to defend a theoretical perspective where changes in educational systems

partly result from endogenous dynamics driven by actors' reactions to others' expected educational behaviours.[4]

As Cherkaoui (*ibid*.: p. 103) himself admits, "the most important" illustration of this idea appears in the account he proposes of the dramatic increase in the fraction of French students who, during the 1950's and 1960s, started to choose the upper secondary track focusing on mathematics and physics over that focusing on humanities (literature and classical languages), in particular among those who wanted to pursue an education in humanities in the most elitist French institutions (*ibid*.: chapter 5). To explain this paradox, Cherkaoui postulates that, when choosing a track, students not only roughly estimate its expected benefits on the job market, but also its selectivity and prestige. To assess the latter, Cherkaoui notes, it is likely that students look at the social composition of the track, and one followed by students regarded as low performers would send a negative signal in terms of quality. As, in the sixties, lower socio-economic groups started to invest in upper secondary school, and in particular in the humanities track, it is likely, Cherkaoui suggests, that students with high educational aspirations started to consider this track as less selective and move to tracks that were believed still to be protected against social devaluation. If so, a dynamic reaction chain can be triggered and a progressive shift from the humanities track to the scientific one may appear. Thus, Cherkaoui proposed an endogenous theory of educational changes that was perfectly in line with the developing areas of formal models focusing on "tipping" points (see, Schelling, 1971) and informational "cascades" (see Granovetter, 1973; Granovetter and Song, 1988) in which dynamic choice interdependence, and its macroscopic unintended consequences, was the crucial dimension.

To conclude, let us stress that these analyses placed Cherkaoui among the leading figures of the sociology of education in France. An indication of this standing is that, in 1986, Cherkaoui wrote the first *Que sais-je* devoted to the sociology of education. By 2010, this short book had reached its eighth edition (see Cherkaoui, 1986).[5]

SOCIAL MOBILITY

Since the first modern large-scale surveys on social mobility (see, for instance, Glass, 1954), educational systems, social inequality in educational outcomes, and occupational mobility are phenomena that sociologists have tended to study in tandem. When looking at the scientific paths taken by students of social stratification, it is frequent indeed to observe a constant back-and-forth between analyses of educational achievement on one hand and research more focused on social mobility on the other (see, for instance, Goldthorpe, 2000).

Cherkaoui is not an exception. His interest in the connection between education and social mobility first appeared at the same time he published his first book on inequalities in educational attainment (see previous section), in particular in a historical piece devoted to a marginal figure of the Durkheimian school, namely Paul Lapie (see Cherkaoui, 1979b [1982]). In this article, Cherkaoui shows that Lapie, to answer the question as to whether formal education was responsible for declining social immobility among agricultural classes' offspring at the end of the nineteenth century in France, constructed and published, in 1904, the first modern inter-generational and career mobility tables in social sciences, very likely without being aware of similar tables published in *Biometrika*, in the same year, by the British statistician Karl Pearson (on this point, see Héran, 2006).

It was, however, toward the end of the eighties, that Cherkaoui's attention started to be directed to social mobility more specifically. In this regard, it will be helpful to note first his analysis of the consequences of actors' perceived status mobility on their perception of what they regard as the status they should deserve (see Cherkaoui, 1988). Data for this analysis came from a French survey on the perception of social inequalities and justice to which Cherkaoui was a leading contributor (*ibid*., 229)[6].

Cherkaoui's analysis ended up with two main sets of results. First, he found that Parisians regarded their social status (self-reported on a 0-lowest–9-highest scale) as more intermediate than it really was and that, based on this perception, they tend to regard their status at best as "fair" but most of time "less than fair" (in any case, never "more than

fair") (*ibid*.: figures 1 and 2). As subsequent research has shown, similar distortions in status perceptions are regularly observed and explanations are increasingly formulated in terms of actors' local interactions (see, for instance, Fararo and Kosaka, 2006). Second, when studying the impact of perceived status mobility on the status respondents wished to have (status mobility is measured in terms of one's perceived status on entry to the job market and ten years later), Cherkaoui found a curvilinear relation. In particular, actors perceiving themselves as downwardly mobile or immobile have, respectively, the strongest and the weakest aspirations for high social status. Between the two extremes, the desire for high social status increases, among those who see themselves as in upward mobility, proportionally to the length of the perceived mobility path (non-linear regression analysis is employed to establish this result).

To appreciate the originality of Cherkaoui's study, it is important to note that, whilst, nowadays, a large literature on the perception of income inequality exists (see, for instance, Niehues, 2014); and extensive international databases are available to study subjective social mobility and its consequences (see, for instance, Kelley and Kelley, 2009), it is common to use subjective status to predict a variety of social and mental outcomes (see, for instance, Singh-Manoux *et al*., 2003). Quantitative, survey-based research on the subjective side of social structure was only in its infancy when Cherkaoui collected and analysed data in 1977 on a sample of Parisians (for another interesting example, see Laumann, 1965).

Cherkaoui's (1995) second main contribution to social mobility research also deals with the consequences of occupational changes but the *explanandum* is not desired status but political behaviour, namely voting. Compared to the previous study, a major difference here is that Cherkaoui's analysis is theoretical in that, though based on empirical literature, it aims to build a "taxonomy" of the effects that social mobility can produce on individuals' political preferences. In particular, distinguishing whether the socially mobile actor's class of origin or destination constitutes the benchmark for his or her final political behaviour, and, on the other hand, whether the actor expresses stronger/weaker political preference (compared to the benchmark), Cherkaoui obtains four elementary types of "political conformism". Then, adding a third

dimension, i.e. the specific electoral choice (right/left), he moves from these basic types to eight more complex cases (*ibid*., 177–182).

Again, to appreciate this contribution, it is important to keep in mind that, whilst recent literature focuses more on class differences in political attitudes (see, for an overview, McCall and Manza, 2011), during the 1950s–1980s, the issue of the political consequences of social mobility was widely debated (see, for an overview, Goldthorpe, 1987: chapter. 1; for an empirical analysis in a comparative perspective, see Turner, 1992). Thus, the usefulness of Cherkaoui's "taxonomy" was to provide a general grid providing order in the heterogeneity of "effects" empirically observed at that time.

This theoretically-oriented analysis is also important for an understanding of the development of Cherkaoui's subsequent research on social mobility. It contains indeed a first, synthetic but explicit, expression of Cherkaoui's growing scepticism about the explanatory power of quantitative research in the field of social stratification, which, he claims, over-values data analysis and downplays the formulation of abstract models from which observations can be deduced (*ibid*.: 171–172, 182).

This scepticism is indeed the premise to Cherkaoui's (2003a) most general contribution to social stratification, in which he proposed "to set some landmarks and to show the way" (*ibid*.: 159) as to the possibility of developing a macro-sociological theory of social stratification. To this end, Cherkaoui built on James Coleman's (1987: 163–168) critique of survey-based methods that tend structurally to be unable to take account of interdependence among actors, thus making it impossible really to understand the transition from micro-decisions to new social structures. Cherkaoui argues that, to develop a truly macro-sociological theory of how one moves from one occupational structure to another, the requirement of interdependence should be applied to every pair of levels (of analysis) through which one must go—what he calls the "principle of generalized interdependence". In particular, he stresses the importance of taking into account *at the same time* the interdependence among educational choices, the competition among job seekers, the interactions between organizations and job seekers, as well as the concurrence among organizations themselves to get the best job seekers—and he

acknowledges that social networks are often the mediator of these various forms of interdependence.

Cherkaoui admits that this is a highly demanding requirement, in particular because of the lack of fine-grained data, but he conveys the important message that one should not take data limitations to justify not trying to develop more focused theoretical lenses. It is too early to say if this message was heeded but it seems fair to say that several recent simulation-based studies go in that direction (see DiMaggio and Garip, 2011; Manzo, 2013a; Fountain and Stovel, 2014).

To conclude, let us stress that, similarly to research in the sociology of education, these empirical and theoretical studies, focusing more directly on social mobility, gave Cherkaoui a lasting standing among the leading figures of sociology of stratification in France. As a sign of this, we should note the two long chapters that Cherkaoui devoted to stratification and mobility in the authoritative *Traité de Sociologie* edited by Boudon (in French) in 1992 (see, respectively, Cherkaoui, 1992a, b). Subsequently, there was the special issue "Mobilité sociale. Histoire, outils d'analyse et connaissance de la société française" that Cherkaoui (in collaboration with L.-A. Vallet) edited in 1995 for the *Revue Française de Sociologie* (see Cherkaoui and Vallet, 1995). Both contributions helped to systematize and institutionalize the field of social stratification in France and provided useful overviews and bibliographic resources for many undergraduate and PhD students.

CLASSICS

As acknowledged by Merton (1967), compared to economics and, even more, physics or biology, sociology "is reluctant to abandon a firsthand acquaintance with the classical works of sociology and pre-sociology as an integral part of the experience of the sociologist *qua* sociologist" (*ibid*.: 30). Mohamed Cherkaoui is a perfect example of this: a dialogue with many of the founding fathers of social sciences such as Durkheim, Weber, Tocqueville or Hobbes, to name a few, is present throughout his publications. Merton also noted, however, that the intimate connection that sociology has with the classics, takes two typical forms, namely "the anemic practices of mere commentary or banalization", which he

regards as "deplorably useless", and "the active practice of following up and developing the theoretical leads of significant predecessors", which Merton considers "wonderfully useful" (*ibid.*: p. 30). Without a doubt, Cherkaoui's relationships with classical sociologists follow the latter path. It is a good illustration of Calvino's (1991 [1999, p. 8]) view according to which "the person who derives maximum benefit from a reading of the classics is the one who skilfully alternates classic readings with calibrated doses of contemporary material". In a word, Cherkaoui is a "presentist": he studied the classics with an eye to their utility for solving contemporary problems (on the distinction between "presentism" and "historicism", which I use here in a neutral way, and academic conflicts related to it, see Borlandi's [2000] profound analysis).

This approach was visible from Cherkaoui's early work in the sociology of education field, which, to a large extent, is consubstantial to Cherkaoui's first dialogue with Durkheim. In particular, this is a deep analysis of Durkheim's *Moral Education* and *The Evolution of Educational Thought* that led Cherkaoui to build his argument on the "relative autonomy" of educational systems, which, as we have seen, is one of the main ideas driving his own study of the transformations of the French educational system in the period after World War II (see Cherkaoui, 1978 [2008: chapter 6]). Among the internal forces creating this autonomy, Cherkaoui argues, conflicts among social groups animated by varying ideological visions of what should be taught at school are a crucial source, a point raised by Durkheim to explain the long-standing domination of the humanities and classical languages in many educational and historical settings (see Cherkaoui, 1981b). By highlighting the conflictual dimension of Durkheim's sociology of education—an aspect that also led Cherkaoui to compare Durkheim to the British sociologist Basil Bernstein (see Cherkaoui, 1977)—and Cherkaoui was well aware that he was proposing a controversial reading of the French sociologist.[7]

This ambition, i.e. "suggesting a non-conformist interpretation of Durkheim's sociological theory" became the central theme of Cherkaoui's first book-length, comprehensive treatment of this founding figure of French sociology (see Cherkaoui, 1998 [2008, p. 2]). The bulk of this interpretation mainly concerns two points: first, Durkheim's

conception of the relationships between the micro- and macro-level of analysis; 2) Durkheim's conception of scientific explanation (*ibid*.: chapters 1 and 3, respectively).

As to the first point, Cherkaoui shows that several ways of relating the macro- and micro-levels are in fact present in Durkheim's analyses of specific social phenomena (for the latest formulation of this argument, see Cherkaoui, 2011a). Among them, the least well recognized by commentators is the upward transition from the micro to the macro, a transition that Durkheim relates to the existence of interdependence structures among individuals. It is, Cherkaoui points out, precisely this aspect that allows Durkheim to consider new macro-structures as "emergent" phenomena, meaning facts containing new properties compared to those of the individuals who helped to generate them.[8]

If this interpretation was intended to contribute "to the disappearance of the myth of Durkheimian holism" (Cherkaoui, 1998 [2008, p. 18]), as to the second point, i.e. Durkheim's explanatory mode, Cherkaoui also aim to counter frequent textbook readings, in particular those seeing Durkheim as exclusively interested in finding nomic regularities among aggregates (most of the time in terms of variable correlations such as in *Suicide*). Cherkaoui's thesis is that, in fact, "to explain is considerably more for the founder of French sociology. It is to deduce the laws and complex structures from basic propositions that constitute the kernel of the theory" (*ibid*.: 84). According to Durkheim, and contrary to Comte, Cherkaoui notes, these "basic propositions" should formulate hypotheses on the "generative facts", "generative mechanisms", "generative causes", or "modes of production of phenomena": terms, as Cherkaoui stresses, that interchangeably recur in Durkheim's writings. By inspecting Durkheim's masterworks like *The Division of Labour in Society* and *Suicide*, Cherkaoui shows that the proposed "mechanisms" invariably refers to complex dynamic bundles of structural elements, individual motives, and interdependence structures.

Overall, by focusing on the micro-level and interactionist component of Durkheim's analyses, Cherkaoui contributed to the accumulation of evidence in favour of the micro-foundationist and relational interpretations of Durkheim's works (see, in particular, Bearman, 1991; Berk, 2006; Boudon, 1995; Collins, 2004: 32–40). Drawing attention to

the link between interdependence and emergence, Cherkaoui anticipated recent readings that, independently, also regarded Durkheim as one of the first "emergentists" (see Sawyer, 2005: chapter 6). Finally, inhighlighting the "mechanismic" components of Durkheim's explanatory mode, Cherkaoui makes us aware of a precursor of the concept of mechanism-based explanations, around which a large literature has developed over the last two decades (see next section).

A similar focus on the sociological classics also animates Cherkaoui's (2006a [2007a]) study of Weber's works. The focus here is the concept of "unanticipated consequences of purposive social action", to take the title of the often-cited paper Merton published in *American Sociological Review* in 1936. Relying on extensive analyses of Weber's writings and correspondence, Cherkaoui sets out to demonstrate how systematically Weber employs this concept in order to explain specific aspects of social reality. As noted by Turner (2007: 178), Cherkaoui's purpose is less historical than substantive, however, in that his ultimate goal is "to expand the concept, systematize it, and thus make it more valuable for sociology". This is clear from the beginning of Cherkaoui's book, where he remarks that his desire to go back to Weber's writings became an urgent one following a period interacting with physicists and economists working in the field of complex systems and raising questions about the emergence of seemingly unpredictable systemic patterns (Cherkaoui, 2006 [2007: 2])[9].

From this perspective, Cherkaoui first retraces the long history of the concept of unintended concequences before Weber, and, then, after commenting on the variety of expressions Weber employs to label this phenomenon—that include "paradox of consequences" (which appears in Cherkaoui's book title)—, starts travelling across an impressive number of Weber's texts and explores how Weber uses, rather than discusses, the concept. The result is the discovery that Weber systematically pointed out five sources of a mismatch between actors' intentions and systemic consequences. According to Cherkaoui's reconstruction, these sources are: 1) mistakes or imperfect information creating gaps between the actor's goal and means (in other words, limitations of instrumental rationality); 2) the interdependence among individual actions, which is defined as actors' tendency to take into account what they believe other actors'

strategies are; 3) spillover effects, meaning the consequences of actions that unintentionally transfer from one domain to another; 4) conflicts among the internal logic of heterogeneous social domains; 5) conflicts among social domains in which different values dominates. In particular, Cherkaoui suggests that these sources are only elementary types that Weber often uses in complex combinations to explain this or that enigmatic phenomenon.

As noted by Kaelber (2009: 1548, 1549), *Good intentions* "succeeds in presenting Weber's thought and extending it to a variety of Weberian themes in innovative ways". For contemporary research, part of this novelty resides in Cherkaoui's capacity to elaborate on the sources Weber indicated as generative of unintended consequences and cast them, as Turner (2007: 178) noted, in terms of "what, in the currently fashionable parlance, are known as mechanisms" (see next section). Cherkaoui's study of Weber's "paradox of consequences" is also important for French sociology. It is indeed well-known that the reception of Weber in France was, at the very least, non-linear (see Borlandi, 1992). Cherkaoui's book helps to correct this trend.[10]

Finally, readers should be aware that Durkheim and Weber are not the only figures in the classic social sciences to whom Cherkaoui devoted systematic analysis. In this respect, let us first mention his study of Tocqueville whose main goal was to correct the common interpretation that Tocqueville saw the French revolution as a result of relative deprivation mechanisms (see Cherkaoui, 2003c [2005: chapter 1])[11]. Contrary to this view, while Cherkaoui acknowledges that Tocqueville implicitly refers to social comparisons based on reference groups when reflecting on the connection between inequality and feelings of justice in *De la démocratie en Amérique,* he shows that this element is absent from Tocqueville's explanation of why discontent leading to protests first appeared in those French regions where feudal institutions were declining. In this case, Cherkaoui argues, Tocqueville only refers to intra-individual comparisons among different periods of time (at the actor- and system-level) and incorporates them within a far larger explanatory scheme in which different hypotheses (at several levels of analysis) are made with reference to different social groups (in particular, the better-conditions-better-expectations psychological driver only holds for economic elites). Thus, in line

with other analyses stressing Tocqueville's modernity in terms of explanatory style (see, Edling and Hedström, 2009; Elster, 2009, in particular chs. 7 and 9), Cherkaoui contributed to the recognition of Tocqueville's role within the emergence of modern social science (see also Boudon, 2006). On the other hand, distinguishing between inter-temporal (and intra-individual) and inter-group comparisons, he helps to clarify different comparison structures to which the concept of relative deprivation may refer (on this point, and the related difference between psychologists and sociologists, see Pettigrew, 2015).

To conclude, let us mention Cherkaoui's (2010) revisitation of Hobbes, which is also animated by the conviction that, similarly to Tocqueville, the contribution of the British philosopher is still valuable not only for political scientists and philosophers but also for sociologists. With this aim, Cherkaoui uses a methodological perspective to analyse *Leviathan*, and argues that Hobbes devised a method based on developing fictional scenarios—about actors' fundamental motives, heterogeneity, and interactions—, which, by thought experiments, he deductively studied to derive societal consequences about social order and political systems. Thus, according to Cherkaoui, Hobbes, although he was not using formalization, anticipated the principles of much contemporary research using simulation for the study of models of social mechanisms. For those who might be tempted to consider this interpretation as embarrassingly anachronistic, I would refer them to Deffuant *et al*'s (2013) article on *The Leviathan model*, in which opinion dynamics are modelled by posing that actors are driven by a concern for self-esteem and it is shown, by simulation, how a variety of social hierarchies can emerge from this simple hypothesis.[12]

EPISTEMOLOGY

As previous sections have shown, Cherkaoui's research in sociology of education, social stratification, and history of sociological theory was consubstantial to epistemological reflection. Such foundational issues as explanatory modes, potentialities and limitations of data analysis, variety of action theories, or the relationships between levels of analysis, underpinned Cherkaoui's intellectual work since the 1970s. In a paper

that went unnoticed, he also overtly rejected the idea that epistemological reflection can be separated from practical problems arising during empirical research (see Cherkaoui, 1990).

Devoting a section of this chapter to Cherkaoui's epistemological contribution does not seem completely inappropriate, however. When considered as whole, indeed, his work clearly entered a reflexive stage, in particular around 2000, during which epistemological statements were formulated more explicitly and autonomously than before. In particular, his views on three fundamental topics—namely, explanation, action theory, and the micro–macro problem—deserve special attention.[13]

As to explanation, Cherkaoui makes explicit his rejection of the idea that one can explain through correlating variables (see Cherkaoui, 1997a [2005: 37]; 2000 [2005: 98–103]) and vigorously claims the superiority of explaining by postulating mechanisms. By "mechanisms", he means "a group of hypotheses that deal with individuals, their interactions and the social context in which they take place", a bundle of elements that is assumed to be capable of dynamically reproducing the phenomenon under scrutiny. Thus, a mechanism is not an "intermediate variable". A mechanism offers a response to the "why" of a given observed linkage. That is why it is difficult to study mechanisms by means of multivariate statistics (Cherkaoui, 1997a [2005: 49, 54]).

Given this definition, Cherkaoui provides two important qualifications. First, the requirement of formulating hypotheses at the actor-level should not be interpreted in a strict sense. Sometimes, he notes, it is not possible to formulate fully specified actor-level statements but, even in these cases, thinking in terms of mechanisms relying on macro-level variables is still preferable (see Cherkaoui, 1997a [2005: 50, 53–54]; 2000 [2005: 106]). Second, when actor-level hypotheses can be formulated, postulating rational actors should not be considered as an essential requirement (Cherkaoui, 1997a [2005: 54–55]). Thus, compared to other understandings of mechanism-based explanations, including that of Boudon, which was one of the inspirational sources for Cherkaoui, the latter's perspective is more flexible and open to the practical impossibility of always being able to fulfil certain methodological requirements.

Cherkaoui's meta-reflection on the concept of mechanism also has an historical interest. Indeed, apart from reconstructing the recent genealogy of the concept through the works of Herbert Simon, Raymond Boudon and Thomas Fararo, Cherkaoui's approach also reminds us of the older origin of the mechanismic perspectives of biologists at the end of the nineteenth century. In particular, he stresses the role played by the French physiologist Claude Bernard, one of the founders of modern medicine (see Cherkaoui, 1997a [2005: 40–43]; 2000 [2005: 92–94]). This is a notable detail in that contemporary philosophical discussions of mechanism-based explanations, which in turn inform debates in sociology (for a recent overview, see Manzo, 2014), are based on more case studies taken from biology than physics (see Machamer *et al*., 2000). It should be noted, however, that Cherkaoui's historically-informed account of mechanism-based thinking does not end up by denying recent progress. He claims, indeed: "In fact the codification of this methodology as well as the degree of formalization achieved nowadays were unknown even a few decades ago" (Cherkaoui, 1997a [2005: 35]).This differentiates Cherkaoui's view from those of other authors, even among the precursors of the mechanismic movement (see Boudon, 2012), who apparently were less willing to acknowledge recent novelty (for a reply to this objection, see Manzo, 2012).

As to action theory, and, more particularly, rational choice theory, Cherkaoui formulates his view in the clearest way while providing a "spirited critique of James Coleman's position" (Kaelber, 2009: 1548), in particular with respect to Coleman's reading of Weber's *Protestant Ethic* and analysis of the emergence of norms (see Cherkaoui, 2003b [2005: chapter 3])[14]. His main point is clearly started from the beginning: "rational choice theorists would be misleading themselves and us if they presented their theory as a theoretical system capable of replacing all sociological tradition when in fact it is only one component of that very tradition" (*ibid*.: 60). In fact, Cherkaoui's target is a specific version of rational-choice theory, namely expected utility theory. His aim is not to deny the relevance of this narrow version—for instance, by discussing elsewhere Coleman's model of panic, Cherkaoui (2000 [2005: 104–106]) recognizes its capacity to provide final explanation. It is rather a problem of scope that Cherkaoui wants to highlight. In particular, he

points out two limitations. First, expected utility theory is "too limited to account for phenomena that partially involves beliefs" (see Cherkaoui, 2003b [2005: 62]). By beliefs, Cherkaoui in fact has in mind normative beliefs, i.e. beliefs actors to subscribe irrespective of the consequences actions based on them may generate (*ibid*.: 70–71). In this respect, he defends the heuristic value of the concept of axiological rationality, following in this Boudon's theory of cognitive rationality (for an explicit statement on this point, see Cherkaoui, 2000 [2005: 106–110]). Second, Cherkaoui argues that, while expected utility theory can cast the emergence of single norms in terms of intentional, non-intentional, or centralized responses to negative externalities, this strategy does not work for "constellations of norms understood as a coherent set" (Cherkaoui, 2003b [2005: 82]). In addition to these scope-limitations, Cherkaoui raises the problem of the realism of explanations based on expected utility assumptions, a crucial challenge for rational choice theory more generally (for a recent overview, see Manzo, 2013b).

Finally let us consider the micro–macro problem. Here again Cherkaoui's view is nuanced and balanced. Cherkaoui starts with a definition of the macro-level as an emergent phenomenon resulting from the interdependence of lower-lever units (see Cherkaoui, 2003a [2005: 154, 155]). He clarifies however that "recognizing the existence of levels of reality should not lead us to think there are ontologically distinct entities that correspond to different types of objects" (Cherkaoui, 1997b [2005: 116]). To Cherkaoui, the real problem is to "discover paths connecting" levels. In this respect, through a systematic analysis of a wide range of sociological perspectives, which are organized by Cherkaoui into a three-dimensional typology focusing on the type of patterns the theory seeks to explain, the kind of micro-level assumptions used, and the unit of observation/analysis employed to test the theory—, Cherkaoui reaches the following conclusion: "there can be no single solution to the problem of the links between micro- and macrosociology, any more than there can be a single mode for explaining all phenomena" (*ibid*.: 141).

Thus, as noted by Edling (2008: 141), Cherkaoui "challenge(s) the idea that all macro phenomena can be reduced to micro interaction", and, as Turner (2005: ix) puts it, "For Cherkaoui, we do not have to

choose between either the agency of social actors and the causal efficacy of social structures" (on this point, see also Little 2013). In fact, thinking dynamically seems the only real requirement Cherkaoui recurrently formulates about the micro–macro problem (see, for instance, Cherkaoui, 2003a [2005: 74–75]; Cherkaoui, 2003b [2005: 29–30]).

To conclude, it is worth highlighting that Cherkaoui's reflexive stage focusing on foundational issues such as the concept of mechanism, action theory, and the micro–macro debate, did not go unnoticed. Commenting on *Invisible Codes*, Swedberg (2007: 184) claimed: "Together with *Dissecting the Social: On the Principles of Analytical Sociology* (2005) by Peter Hedström, it constitutes the most important contribution to the theoretical discussion of social mechanisms during the last few years". In a similar vein, Edling (2008: 135) noted: "Cherkaoui adds value to contemporary discussions on the role of rational choice theory and social mechanisms, as these discussions come to light in for instance Goldthorpe's *On Sociology* (2007) and Hedström's *Dissecting the Social* (2005), and I believe that both *Invisible Codes* and *Good Intentions* should be read within this framework".

MOROCCAN SOCIETY

Contrary to what one may have expected, more or less at the same time as *Invisible Codes* was published, in 2005, Cherkaoui, rather than moving towards even more abstract thinking, initiated a new wave of empirical research in a new domain, namely Moroccan society and geopolitics. Thus, he addressed, in turn, such complex topics as fundamentalist Islamism (Cherkaoui, 2006b [2007a: chapter 8]), the historical relationships between Morocco and Western Sahara (Cherkaoui, 2007b), institutional trust in Morocco (Cherkaoui, 2009), and the Moroccan university system (Cherkaoui, 2011b)[15].

Before retracing some of the main ideas and findings of these studies, let us note that Cherkaoui himself admits that, for a long time, "the Maghreb [does] did not fall within the scope of my research topics" (Cherkaoui, 2007b: ix). The reason, he continues, is that "Cultural area sociological studies have never enjoyed a special status within the stratified world of scientific research and, except in a few instances,

it is something more often associated with demanding journalism than fundamental research seeking a solution to sociological puzzles by combining theories and empirical data rather than the pure narration of accumulated facts". It is precisely with this spirit that Cherkaoui undertook the analysis of the Maghreb: all the historical, theoretical, and methodological *savoir-faire* he accumulated over the years, is now applied to this new topic.[16]

Cherkaoui's (2006b [2007a: chapter 8]) analysis of Islamic fundamentalism is a first expression of this ambition. The explanation he outlined indeed originally combines the historical sociology of religion of Mauss, Durkheim, and Weber with Durkheim's concept of "anomie" and "altruistic suicide" as well as more recent rational-choice inspired analyses (where James Coleman is again, both a source of inspiration and a critical target). The diffusion of fundamentalist movements in Morocco and elsewhere is thus conceived at the intersection of complex phenomena referring to several levels of analysis, which, Cherkaoui notes, "are inextricably entangled with each other" (*ibid*.: 173).

In particular, at the "morphological" level, Cherkaoui stresses the existence of secular tensions within Muslim societies, between different interpretations of Islam that are directed to different social groups with different needs and social interests. Recent evolution towards forms of orthodox and puritan Islam must be understood as the encounter between these long-term conflicts and new structural transformations of Muslim societies, in which a growing fraction of the population is becoming urbanized and seeks new sources of values, identity, and social belonging (the "axiological" level). At the "organizational" level, as Islam lacks a legitimate institution that monopolizes the interpretation of religious texts, these structural transformations create the condition for a proliferation of doctrinal centres all competing to provide moral support for those who have difficulties in coping with a quickly changing society. The more the discourse of these doctrinal centres attributes problems to external sources, Cherkaoui notes, the more it is acceptable to all those who feel threatened by the current system, the easier it is for doctrinal groups to affiliate and completely (and impressively quickly) re-socialize their members. It is only within these complex social dynamics, Cherkaoui concludes, that extreme individual

behaviours, like suicide bombers, can be understood. Paradoxically, it is the quest for social integration that, through the re-creation of small groups with strong internal cohesion, can lead an individual into self-harming behaviour. "Suicide is carried out as if it were a duty", Cherkaoui claims (ibid., p. 190).

Thus, Cherkaoui defies recurrent interpretations in terms of cost-benefit calculus of the most shocking expressions of fundamentalist movements (see, for instance, Wintrobe, 2006), and invites us to think of this phenomena within a wider theoretical context taking into account the long-term features of Muslim societies and their internal development, aspects that seem to escape many Western observers.

Similarly stimulating analyses characterize Cherkaoui's (2007b) subsequent study of the relationships between Morocco and Western Sahara. This is a complex geopolitical issue, involving, through a secular process, Western (France and Spain) and Arabic (Algeria) countries. As Cherkaoui admits from the beginning, the analysis is complicated by the fact that the issue is "loaded with so much emotion that I might be accused of bias" (*ibid*.: ix). To deal with this difficulty, Cherkaoui combines ideas and methods from three different, although complementary, disciplines, namely history, international relations, and sociology. As Cherkaoui admits (*ibid*., p. x), the sociological investigation is the most original part of the book.

In this part of the study (forming the second half of the book), Cherkaoui's goal is to demonstrate that, despite historical and political struggles, and external political interventions (analyzed in the first part of the book), there are deep structural reasons to argue that Morocco and Western Sahara are in fact strongly integrated. By integration, Cherkaoui means "the volume, extension, intensity and density of the reciprocal social relationships that the Saharawis entertain with other Moroccans" (*ibid*.: 73). To accumulate evidence in favour of his argument, Cherkaoui first exploits a variety of national-level and local administrative data-sources covering the last four decades and establishes the existence of sizeable improvements in education (*ibid*.: chapters 5 and 6), leaving conditions, measured in terms of poverty rates (*ibid*.: chapter 7), and basic infrastructures (*ibid*.: chapter 9) among inhabitants of Western Sahara. Cherkaoui stresses that these changes

are similar to, if not better than, those observed within other Moroccan regions, and largely due to the massive investmentsby the Moroccan authorities in the Sahara provinces.

But, as Cherkaoui claims repeatedly, in his view, the best indicator of social integration between Sahrawis and other Moroccan will be found elsewhere, namely in marriages (*ibid*.: chapter 8). That is why Cherkaoui had approximately 30.000 marriage contracts collected and manually coded in order to assess "whether people born in the Sahara are becoming increasingly heterogamous or whether they remain strictly homogamous and indeed endogamous as in the past (*ibid*.: 139). Analysis of this data indeed shows that living, among Sahrawis in Sahara, the proportion of those who marry a woman/man residing in "other Saharan regions of southern Morocco" or "all other regions of Morocco" markedly decreased between 1947 and 2007 by 45.3% for men and 39% for women (*ibid*.: 142–143), a decreasing trend observed across all occupational groups (*ibid*.: 149).

Thus, in addition to supporting his chief argument, Cherkaoui, through an original and demanding data collection procedure, indirectly contributes to the international literature by complementing, with more fine-grained and recent information, the rare quantitative studies of homogamy in comparative perspective that contain data on Morocco (see Smits *et al*., 1998, 2003)[17].

Creativity and originality in data collection procedures also characterizes Cherkaoui's (2011) other major contribution to the analysis of Moroccan society, dealing more particularly with the Moroccan university and the social sciences. The first part of this study addresses issues in the sociology of education and science such as the training paths of professors, their career mobility, their prestige, as well as their capacity to form a scientific community. Interview-based and quantitative analysis of survey data that Cherkaoui collected led to the major conclusions that, in the social sciences within Moroccan universities, institutional rules favour hierarchy and status over merit, budgets are sub-optimal, professors' social prestige is very low, and peer-oriented practices are almost inexistent and scarcely valued by professors themselves. Moroccan professors in the social sciences, Cherkaoui concludes, do not form a scientific community *sensu stricto* (*ibid*.: 115).

The second part of Cherkaoui's analysis concerns more specifically the scientific output of Moroccan social science. In this respect, since available citation databases do not reference Moroccan social science journals, and, the percentage of Moroccan social scientists publishing in referenced journals is close to zero, quantifiable information had to be created. Thus, by using and connecting several Moroccan sources, Cherkaoui was able to generate an original database containing approximately 57000 references, covering virtually all the scientific output (articles, books, and other documents) between 1960 and 2006. A very distinctive feature of this database is that it also provides information on the output of non-academics, thus allowing bibliographic comparisons of professors and non-professors in social sciences (*ibid*.: 121–122, 131–135). This aspect led to one of the most original results of Cherkaoui's study. In particular, Cherkaoui discovered that, in Moroccan social science at least, both the form (in terms of asymmetry) of the distribution of publications per author (*ibid*.: chapter 3) and its temporal evolution (*ibid*.: chapter 5) are remarkably similar among academics and authors from outside university.[18] According to Cherkaoui, this is the result of the fact that both academics and non-academics are in fact players in the same intellectual market, namely one where political demands, mass media pressure or general public interest weigh more than their peers' evaluation (*ibid*., pp. 159, 191, 215–216). Cherkaoui submits that this may be a general phenomenon that should be observed elsewhere provided the following structural condition exists: academic gratification is so weak that academics prefer to seek social recognition in non-academic domains[19].

CONCLUSION

A systematic analysis of Cherkaoui's work suggests that his scientific trajectory evolved through five main stages. First, in the 1970s and 1980s, his research was based on refined quantitative methods and deals with educational inequalities in comparative perspective as well as with the specificities of the French educational system and its historical transformation. Second, from the end of the 1980s and until the mid-1990s, Cherkaoui's work extended to social mobility, in particular its subjective

aspect and individual-level consequences. Third, between around the mid-1990s and mid-2000s, Cherkaoui's focus was on the classics, especially Durkheim and Weber. Fourth, during the 2000s, Cherkaoui's epistemological thinking became more autonomous as compared to previous stages. Fifth, starting in the second half of the 2000s, Cherkaoui initiated a new phase of empirical research dealing with several aspects of Moroccan society.

Obviously, this periodization has limitations. It is based on publication dates, which do not necessarily reflect the evolution of Cherkaoui's research interests. Moreover, publication dates themselves suggest overlaps. As I have stressed several times, it is clear, for instance, that the classics, namely Durkheim, occupied Cherkaoui's mind since his early research in sociology of education. Thus, as the title of the present introductory essay suggests, the periodization proposed should be understood as akin to a road map providing readers with some guidelines pointing out major thematic focuses and lines of development in Cherkaoui's works. Let me also acknowledge that the focuses and evolutions I highlighted may be criticized for not doing full justice to all the empirical results and theoretical/ methodological nuances that characterize his scientific work. This certainly results from the specific viewpoint from which I reviewed Cherkaoui's writings, a specificity that is related to my own sociological agenda as well as to my own relationship with Mohamed Cherkaoui. I cannot exclude that the systematic discussions that we have had over the years on this or that aspect of his sociology led me to accentuate one or other dimension of his publications.

But, as Italo Calvino (1991 [1999, p. 5]) noted in his essay on "why read the classics" that I mentioned in the introduction, "no book which discusses another book can ever say more than the original book under discussion". From this point of view, fortunately, I am not alone. Contributors to this *Festschrift* will balance, correct, and complement the possible limitations of the road map I proposed in this introduction. They and I, in the spirit of Calvino's remark, should ultimately be regarded as a humble invitation to read and delve into the details of Cherkaoui's writings more than all we did in the present *liber amicorum*.

To conclude, I would like to devote a few words to its genesis and structure. As to the first aspect, let me recall two facts. First, as

provocatively noted by Tulving (2007, p. 39), "festschrift frequently enough also serves as a convenient place in which those who are invited to contribute find a permanent resting place for their otherwise unpublishable or at least difficult to publish papers." As a consequence of this, others noted, "All too often, the festschrift consists of a disparate and uneven collection of papers on a range of subjects that often only vaguely intersect with the interests of the individual whose work is being honored by the volume." (Nunan, quoted in Tulving, 2007, p. 40). I am not sure I succeeded in avoiding the first shortcoming—even though I rejected some of the paper proposals I initially received and have not published some of the papers that finally reached me. I did my best, however, to escape the second drawback. To this aim, contributors were originally invited (in June 2012) explicitly to write a piece in relation to at least one aspect of Cherkaoui's work. Despite the general, and, after all, weak, nature of my requirement, the emergent result was a pleasant surprise.

The contributions finally published do indeed cover all major thematic areas, methods, and stages of development through which Cherkaoui's work has evolved over the years. From educational inequalities to social mobility and its consequences; from the classics and sociological theory to epistemology, up to more recent topics such as Morocco, social movements, and international relations, all of Cherkaoui's writings were used more or less directly by contributors to think, more or less critical, about their heuristic utility. For this effort, and for patiently accepting my requests for changes and corrections, I would warmly thank all of them.

As to the structure of the book, I attempted to organize the thirty-five contributions by combining two criteria. The first concerns the theory/empirical divide so that the first part of the book (titled "Theories") contains all contributions in which a theoretical orientation was dominant whereas the second part (titled "Social mechanisms") contains contributions in which the empirical orientation is given priority, at least in the sense that a specific class of empirical phenomena constitutes the target of the analysis. Within each part, contributions are organized in sub-sections in such a way that the sections' order reflect the temporal development of Cherkaoui's research interests (within each

sub-section, contributions are listed in alphabetic order—when more than one author is present, the family name of the first author has precedence). This organizing principle also has its limitations, but it is the least arbitrary choice I could make.

My hope is that the present introduction combined with the contributors' essays help readers to better see that when Cherkaoui's work is considered as a whole, it conveys important messages for contemporary sociology. First, Cherkaoui proposes a clear vision of social phenomena as arising from complex systems of interdependence in which the behaviour of low-level entities constantly trigger, often unintentionally, high-level patterns that subsequently feedback into later behaviour. Second, he invites us to go behind empirical description, namely trying to open up "black boxes" that connect several levels of analysis. Third, Cherkaoui enjoins us to employ data analysis and multivariate statistics in a critical way by keeping constantly in mind data limitations and the distinction between description and explanation. Fourth, he exhorts us to maintain a fruitful dialogue with classics, not only from sociology, not merely to recount what they contain, or discover what they were really trying to say, but as a source of intellectual stimulation and emulation. Finally, Cherkaoui is an example of how fruitful a systematic back-and-forth between empirical, historical, and epistemological research can be.

Some observers, in commenting on *Invisible Codes*, claimed that Cherkaoui's work was a sign of the distinctiveness of French sociology compared, for instance, to British sociology, which, in their view, spends most of its energy on language, culture, and post-modernism, often using unclear prose and superficial methods (see Turner, 2005; Webster, 2006). In my humble opinion, this judgment results from a position bias, i.e. the tendency to know more about "in-groups" than "out-groups". Unfortunately, Cherkaoui's sociology is not representative of French sociology. The theoretical and methodological messages his *oeuvre* conveys suggest however the way to make it better.

NOTES

1. As a methodological note, in order better to see the chronological evolution of Cherkaoui's research interests, I will systematically refer to the dates of the original publications in French. However, in order to allow English-speaking readers to check my claims, I will equally systematically give (in brackets) the date of the first English translation, when available. In these cases, references to pages will refer to the English translation.
2. This book originated from Cherkaoui's PhD dissertation (supervised by Raymond Boudon and defended in 1975 at the university of Paris-Sorbonne). Interestingly, the copy of the book stored in the library of the University of Paris-Sorbonne (Serpente) contains a dedication to Jean Stoetzel. It is worth noting that, historically, Jean Stoetzel played a crucial role in introducing quantitative methods in France. He founded, in 1934, one of the major opinion polling institutes in France (i.e. the *Institut Français d'Opinion Publique*, IFOP), as well as, in 1959, the leading French sociological journal (i.e. the *Revue Française de Sociologie*). Significantly, Jean Stoetzel was Raymond Boudon's PhD supervisor (see Girard, 1987).
3. These changes are: (1) the increasing rates of participation at the upper secondary and tertiary levels; (2) the decreasing trend in the age of *baccalauréat* candidates; (3) the increase in the proportion of female students; (4) the loss of attraction of the humanities track at the upper secondary level.
4. For the reactions that the empirical critique Cherkaoui addressed on this basis in particular to Crozier and Bourdieu's explanations of educational expansion (*ibid*.: chapter 2), see Dubar (1984), to whom Cherkaoui (1984) replied.
5. *Que sais-je* is a prestigious series issued by the Presses Universitaires de France that publishes book-length syntheses of a given research field.
6. The survey was based on a non-probabilistic stratified sampling of 1000 men and women from Paris and its suburbs. It was realized in 1977 and benefited from a research grant the GEMASS (the CNRS research unit to which Cherkaoui belonged and that was headed at that time by Raymond Boudon, who founded it in 1971) received, in 1974, from the Commissariat Général au Plan d'Équipement et de la Productivité. Cherkaoui authored, in 1977, the second part of the document reporting on the survey to the Commissariat. A quick look at the questionnaire, annexed to this document, suffices to show that this was an innovative survey, certainly among the first to be entirely devoted to the subjective aspects of inequality structures.
7. He noted: "To summarize my critics: I am supposed to have devised a marxist, tendentious, and paradoxical interpretation of the text, because I placed the category of conflict at the centre of the Durkheimian theory of educational systems" (see Cherkaoui, 1981b: 127).
8. Note that the idea that Durkheim can be read as an "emergentist" already appears in Cherkaoui (1982: 101–102), he refers to an "effet rétro-coercitif" (*constraining feedback*, my translation) to indicate social dynamics in which actions at time t crystallize into aggregates that become constraints for later actions. Cherkaoui claims that Durkheim's definition of institution relies on this circularity.

9. It is important however to keep in mind that Cherkaoui's passion for high-level unintentional consequences is already visible in his early work in the sociology of education. In particular, Cherkaoui (1982: 101–103) contains a long list of references in which the concept is present. The second chapter of *Good intentions* is in fact a detailed development of this list. Thus, although it is true that Cherkaoui's publications on Weber appeared later than those on Durkheim—*Good Intentions* is indeed preceded only by two articles on the German sociologist Cherkaoui published in the early 2000s (see, Cherkaoui, 2003b [2005: chapter 3]; 2004), the perspective from which Cherkaoui navigates through Weber is based on ideas that had driven Cherkaoui's scientific output since the 1970s.
10. *Good Intentions* also raised some objections, in particular by Kaelber (2009: 1549) and McGovern (2010: 1206), who complained about Cherkaoui's tendency to downplay critical secondary literature on Weber and more recent developments on unintentional consequences.
11. Among these inaccurate accounts, Cherkaoui (*ibid*., p. 22) mentions the one Coleman delivered in *Foundations of Social Theory*. Interestingly, Coleman was also chosen as a critical target by Cherkaoui with respect to Coleman's reading of Weber's explanation of the emergence of capitalism (see Cherkaoui, 2003b [2005: chapter 3]).
12. Let me remind the reader that Cherkaoui's interest in sociological theory and the history of social thought also appeared in his collaboration in some sizeable national and international editorial projects, in the form of dictionaries (see Boudon *et al*., 2005) or (monumental) commented collections of texts (see Alexander, Boudon, and Cherkaoui, 1997a, b; Boudon and Cherkaoui, 2000a, b).
13. As an indication of this reflexive, epistemological turn, consider what Cherkaoui (2005, p. 1) writes in the introduction to *Invisible Codes*: "The current work is the result, a very restricted one, of my thinking over the last thirty years. My interest in the problem of generative mechanisms is something I owe to the teaching of Raymond Boudon (…). A large part of my empirical work on stratification and social mobility was worked out in this tradition (Cherkaoui 1979, 1982, 1988, 1992, 1995). The more theoretical and historical aspects of this work offered here to the reader share a similar ambition."
14. To explain Cherkaoui's focus on Coleman, it should be recalled that his contribution was part of a special issue that the *Revue Française de Sociologie* (2003, 44, 2) decided to devote to James Coleman's *Foundations* in order to make French sociology aware of his important work. Following the order of this issue's table of contents, Philippe Steiner, Peter Abell, Olivier Favereau, Jon Elster, Emmanuel Lazega, Alban Bouvier, Siegwart Lindenberg, Steve Lukes, and Raymond Boudon also contributed to this volume. To a large extent, rational choice theory is discussed in every contribution.
15. Other works are in progress. Cherkaoui briefly describes them in the interview closing this *Festschrift* (see Cherkaoui, 2015, present volume).
16. In passing, it should be noted that a short article related to Morocco, namely dealing with marriage homogamy in Casablanca, had in fact already been published by Cherkaoui in the 1990s (see Cherkaoui, 1994).

17. The political implications of this analysis (coupled with statements on Algeria Cherkaoui made in the first part of the book) quickly led to Cherkaoui's study being translated in French, Spanish and Arabic, and generated a considerable mass-media coverage in several countries. In this regard, among the most significant interventions, consider the "opinions" Cherkaoui published in *The Washington Times* (see Cherkaoui, 2007a, b).
18. In passing, note that Cherkaoui's interest in asymmetric distributions, their description and generative models goes back to his early research on educational inequalities (see, for instance, Cherkaoui's [1982: 154–156] and his remarks on the age distribution of French *baccalauréat* candidates).
19. Cherkaoui's severe critical analysis of the Moroccan university had such a profound impact on the Moroccan political milieu that the King himself took inspiration from Cherkaoui's conclusions to launch a new reform initiative on the Moroccan educational system (see Cherkaoui, 2013).

REFERENCES

Cherkaoui's Works

Alexander, J., Boudon, R., Cherkaoui, M. (eds) (1997a) *The Classical Tradition in Sociology: The American Tradition* (four volumes). London: Sage.

——— (eds) (1997b) *The Classical Tradition in Sociology: The European Tradition* (four volumes). London: Sage.

Boudon, R., Borlandi, M., Cherkaoui, M., Valade, B. (eds) (2005) *Dictionnaire de la pensée sociologique*. Paris: Presses Universitaires de France.

Boudon, R., Cherkaoui, M. (eds) (2000a) *Central Currents in Social Theory: The Roots of Sociological Theory 1700–1920* (four volumes). London: Sage.

——— (eds) (2000b) *Central Currents in Social Theory: Contemporary Sociological Theory 1920–2000* (four volumes). London: Sage.

Cherkaoui, M. (1976) "Transformation to normality: Variations of the power transformation. parameters under various sociological conditions", *Quality and Quantity*, 10: 179–184

——— (1977) "Bernstein and Durkheim: Two theories of change in educational systems", *Harvard Educational Review*, 4(47): 556–564.

——— (1978) "Système social et savoir scolaire: les enjeux politiques de la distribution des connaissances", *Revue Française de Science Politique*, 28(2): 313–348.

——— (1979a) *Les paradoxes de la réussite scolaire. Sociologie comparée des systèmes d'enseignement*. Paris: Presses Universitaires de France.

——— (1979b) "Les effets sociaux de l'école selon Paul Lapie", *Revue Française de Sociologie*, 1(20): 239–255 [English translation as: "Education and social Mobility: Paul Lapie's pathbreaking research". In Besnard, P., *The Durkheimians: Constructing the Sociological Domain*. Cambridge: Cambridge University Press, 1982: 217–231].

——— (1981a) "Changement social et anomie: essai de formalisation de la théorie durkheimienne", *Archives Européennes de Sociologie*, 22(1): 3–39.

——— (1981b) "Consensus or conflict? Return to Durkheim's proteiform theory", *Theory and Society*, 10: 127–138.
——— (1982) *Les changements du système éducatif en France 1950–1980*. Paris: Presses Universitaires de France.
——— (1984) "Réponse à la note critique de Claude Dubar", *Revue Française de Sociologie*, 25(2): 327–329.
——— (1986 [2010, 8th edition]) *Sociologie de l'éducation*. Paris: Presses Universitaires de France.
——— (1988) "Mobilité sociale et équité", *Revue Française de Sociologie*, 29(2): 227–245.
——— (1990) "Du pluralisme méthodologique: pour éviter les pièges du dogmatisme et de l'anarchie", *Revue Européenne des Sciences Sociales*, 28(88): 97–111.
——— (1992a) "Stratification", in R. Boudon (ed.), *Traité de Sociologie*. Paris: Presses Universitaires de France, 97–152.
——— (1992b) "Mobilité", in R. Boudon (ed.), *Traité de Sociologie*. Paris: Presses Universitaires de France: 153–193
——— (1994) "Les deux modèles dominants du marché matrimonial au Maroc", foreword to M. Aboumalek (1994) *Qui épouse qui? Le mariage en milieu urbain*. Ed. Afrique Orient: 9–18.
——— (1997a) "La théorie de l'explication chez Durkheim: modèle déductif-nomologique et modèle des mécanismes générateurs", in C.-H. Cuin (ed.), *Durkheim d'un siècle à l'autre : Lectures actuelles des Règles de la méthode sociologique*. Paris: Presses Universitaires de France [English adaptation published as "Composition effect and generative mechanisms. An interpretation of Durkheim's theory of integration and regulation", in Cherkaoui [2005: chapter 2])
——— (1997b) "Le réel et ses niveaux : peut-on toujours fonder la macrologie sur la micrologie?" *Revue Française de Sociologie*, 38(3): 497–524 [English translation, as "The multiple levels of reality: Is the macro always based on the micro", in Cherkaoui [2005: chapter 5])
——— (1998) *Naissance d'une science sociale. La sociologie selon Durkheim*. Geneva/Paris: Droz (English translation: see Cherkaoui [2008])
——— (2000) "La stratégie des mécanismes générateurs comme logique de l'explication", in J. Baechler and F. Chazel (eds), *L'acteur et ses raisons. Hommages à Raymond Boudon*. Paris: Presses Universitaires de France [English translation in Cherkaoui [2005: chapter 4]).
——— (2001) "Relative deprivation", in N.J. Smelser and P.B. Baltes (eds), *The International Encyclopaedia of Social and Behavioral Sciences*. Amsterdam: Elsevier.
——— (2003a) "The individual and the collective," *European Review*, 11(4): 489–504 (reprinted in Cherkaoui [2005: chapter 6] as "Notes on the syntaxes of stratification theories").
——— (2003b) "Les transitions micro–macro. Apports et limites de la théorie du choix rationnel dans les *Foundations of Social Theory*", *Revue Française de Sociologie*, 43 [English translation in Cherkaoui [2005: chapter 3])
——— (2003c) "L'Etat et la Révolution. Logique du pouvoir monopoliste et mécanismes sociaux dans L'Ancien Régime de Tocqueville", *Tocqueville Review/Revue*

Tocqueville, 1(24) (English translation in Cherkaoui [2005: chapter 1])
——— (2004) "Le réel et le rationnel. Rationalité et conséquences inattendues chez Max Weber", *Revue Européenne des Sciences Sociales*, XLII-129: 73–86.
——— (2005) *Invisible Codes*. Oxford: Bardwell Press.
——— (2006a) *Le Paradoxe des conséquences. Essai sur une théorie wébérienne des effets inattendus et non voulus des actions*. Geneva/Paris: Droz (English translation: see Cherkaoui [2007a]).
——— (2006b) "Le fondamentalisme islamique. Esquisse d'une interprétation", *Commentaire*, 114: 349–361.
——— (2007a) *Good Intentions. Max Weber and the Paradox of Unintended Consequences*. Oxford: Bardwell Press.
——— (2007b) *Morocco and the Sahara. Social Bonds and Geopolotical Issues*. Oxford: Bardwell Press, 2007
——— (2007c) "The Mediterranean union. A tool of will and representation", *The Washington Times*, Monday, December 3, 2007, A15.
——— (2007d) "Western Sahara. The missing dimension", *The Washington Times*, Monday, December 3, 2007, A19.
——— (2008) *Durkheim and the Puzzle of Social Complexity*. Oxford: Bardwell Press.
——— (2009) *La confiance dans les institutions au Maroc*. Rabat: Institut Royal des études stratégique (IRES).
——— (2010) T"homas Hobbes: On generating social order", in Christofer Edling and Jens Rydgren (eds), *Sociological Insights of Great Thinkers. Sociology through Literature, Philosophy, and Science*. Oxford: Praeger: chapter 32.
——— (2011a) "La complexité du social et la méthode sociologique selon Durkeim", in R. Boudon (ed.), *Durkheim fut-il durkheimien?* Paris: Armand Colin: 81–98.
——— (2011b) *Crise de l'université. Le nouvel esprit académique et la sécularisation de la production intellectuelle*. Geneva/Paris: Droz.
——— (2013) "La réforme doit être totale et visionnaire", Atlasinfo (Interview by *Mostafa Bouaziz, Reda Mouhsine et Abdellah Tourabi*) (accessed on 2 April at: http://zamane.ma/fr/interview-la-reforme-doit-etre-totale-et-visionnaire/).
——— (2015) "Cross-cutting fields and social circles: A conversation with Mohamed Cherkaoui", in Manzo, G. (ed.) *Theories and Social Mechanisms: Essays in honor of Mohamed Cherkaoui*. Oxford: Bardwell Press.
Cherkaoui, M., Lindsey, J. (1977) "Problèmes de mesure des classes sociales: des indices du status aux modèles d'analyse des rapports de classe", *Revue Française de Sociologie*, 18(2): 233–270
Cherkaoui, M., Vallet, L-A. (1995) "Presentation", *Revue Française de Sociologie*, 36(1): 1–3 (special issue on "Mobilité sociale: Historie, outils d'analyse et connaissance de la société française").
Lindsey, J., Cherkaoui, M. (1975) "Some aspects of social class differences in achievements among 13-year-olds", *Comparative Education*, 3(11): 247–260.

Further works by Mohamed Cherkaoui not quoted in the present article

Borlandi, M., Cherkaoui, M. (eds) (2000) *Le suicide. Un siècle après*. Paris: Presses Universitaires de France.

Boudon R., Besnard, P., Cherkaoui, M., Lécuyer, B.-P. (eds) (1989) *Dictionnaire de sociologie*. Paris: Larousse.

Boudon, R., Bulle, N., Cherkaoui, M. (2001) *Ecole et société*. Paris: Presses Universitaires de France.

Cherkaoui, M. (2012) "Stratégies de conflit et ellipse des crise", *Annuaire Marocain de la Stratégie et des Relations Internationales*, 2012.

——— (2014) "Incertitude et ellipse des crises", in D. Vrancken, *Penser l'incertitude*. Laval: Presses Universitaires de Laval.

——— (2014) *L'arc tendu de l'insurrection. Esquisse d'une explication des révoltes arabes*. Beirut: Presses de la Fondation de la Pensée Arabe.

——— (2015) "Désenchantement du monde, dépeuplement du ciel: étude comparée des sphères normatives", in M. Janjar, *Du Maghreb et d'ailleurs. Mélanges en hommage à Al Aydi*. Casablanca: Editions Okad.

Comments on Cherkaoui's work

Backhaus, J. (2014) "*Mohamed Cherkaoui: Morocco and the Sahara, Social Bonds and Geopolitical Issues*", *European Journal of Law and Economics*, 37(2): 355 (not quoted in this chapter).

Dubar, C. (1984) "Sur deux ouvrages de sociologie de l'éducation", *Revue Française de Sociologie*, 25(1): 151–158.

Edling, C. (2008) "Mohamed Cherkaoui: *Invisible Codes: Essays on Generative Mechanisms* Mohamed Cherkaoui: *Good Intentions: Max Weber and the Paradox of Unintended Consequences*", *European Sociological Review*, 24(1): 135–139.

Kaelber, L. (2009) "*Good Intentions: Max Weber and the Paradox of Unintended Consequences*", *American Journal of Sociology*, 114, 1547–49

Little D. (2013). "Mohamed Cherkaoui on the micro–macro debate. Understanding society" (http://understandingsociety.blogspot.fr/2013/04/mohamed-cherkaoui-on-micro-macro-debate.html) (accessed on March, 31 2015)

McGovern, P. (2010) "Book review: Mohamed Cherkaoui, *Good intentions: Max Weber and the Paradox of Unintended Consequences*", Oxford: The Bardwell Press, 2007, *Sociology*, 44, 6, 1205–1206.

Swedberg, R. (2007) "Mohamed Cherkaoui, *Invisible Codes: Essays on Generative Mechanisms*", *Contemporary Sociology*, 36(2): 184–85.

Turner, B. S. (2005) "Foreword: Visible forms and invisible codes: Description, interpretation, and explanation", in M. Cherkaoui (2005), *Invisible Codes. Essays on Generative Mechanisms*. Oxford: Bardwell Press.

Turner, S. P. (2007) "Mohamed Cherkaoui: *Good Intentions. Max Weber and the Paradox of Unintended Consequences*", *Contemporary Sociology*, 37(2): 178–179.

Webster, F. (2006) "Invisible Codes", *The Times Higher Education*, 2 June 2006 (http://www.timeshighereducation.co.uk/books/at-the-core-of-living/203542.article) (accessed on April, 2 2015)

General References

Calvino, I. (1991 [1999]) *Why Read the Classics?* New York: Vintage Books.

Chan, T. W., Goldthorpe, J. H. (2007) "Class and status: the conceptual distinction and its empirical relevance", *American Sociological Review*, 72: 512–532.

Singh-Manoux, A., Adler, N. E., Marmot, M. M. (2003) "Subjective social status: its determinants and its association with measures of ill-health in the Whitehall II study", *Social Science & Medicine*, 56(6): 1321–1333.

Bearman, P. S. (1991) "The social structure of suicide", *Sociological Forum*, 6(3): 501–524.

Berk, B. B. (2006) "Macro–micro relationships in Durkheim's analysis of egoistic suicide", *Sociological Theory*, 24(1): 58–80

Borlandi, M. (1992) "Max Weber and/in French sociology: a book review", *Revue Européenne des Sciences Sociales*, 30(93): 103–121

——— (2000) "La querelles des historiens et des présentistes", in J. Baechler and F. Chazel (eds), *L'acteur et ses raisons. Hommages à Raymond Boudon*. Paris: Presses Universitaires de France, 88–112.

Boudon, R. (1974) *Education, Opportunity, and Social Inequality*. New York: John Wiley & Sons, Inc.

——— (1995) "Should one still read Durkheim's rules after one hundred years?" *Revue Suisse de Sociologie,* 21(3), 559–573.

——— (2006) *Tocqueville for Today*. Oxford: Bardwell Press.

——— (2012) "'Analytical sociology' and the explanation of beliefs", *European Journal of Social Sciences*, 50(2): 7–35.

Bourdieu, P., Passeron, J-C. (1970). *La reproduction. Éléments pour une théorie du système d'enseignement*. Paris: Editions de Minuit.

Coleman, J. S. (1966) *Equality of Educational Opportunity* (accessible at: http://doi.org/10.3886/ICPSR06389.v3)

——— (1987) "Microfoundations and macrosocial behavior", in J. C. Alexander, B. Giesen, R. Münch, N. Smelser (eds), *The Micro–Macro Link*. Berkeley/Los Angeles, CA: University of California Press.

Collins, R. (2004) *Interaction Ritual Chains*. Princeton and Oxford: Princeton University Press.

Deffuant, G., Carletti, T., Huet, S. (2013) "The leviathan model: Absolute dominance, generalised distrust, small worlds and other patterns emerging from combining vanity with opinion propagation", *Journal of Artificial Societies and Social Simulation* 16(1): 5: http://jasss.soc.surrey.ac.uk/16/1/5.html.

DiMaggio, P. and Garip, F. (2011) "How network externalities can exacerbate intergroup inequality", *American Journal of Sociology*, 116(6): 1887–1933.

Edling, C. R. (2002) "Mathematics in sociology", *Annual Review of Sociology*, 28: 197–220.

Edling, C. and Hedström, P. (2009) "Tocqueville and analytical sociology", in C. Mohamed and P. Hamilton (eds), *Raymond Boudon: A Life in Sociology. Essays in Honour of Raymond Boudon*. Oxford: Bardwell Press.

Elster, J. (2009) *Alexis de Tocqueville: The First Social Scientist*. New York: Cambridge University Press.

Fararo, T. J., Kosaka, K. (2006) *Generating Images of Social Stratification: A Formal Theory*. Dordrecht: Kluwer Academic Press.

Fountain, C., Stovel, K. (2014) "Turbulent careers: Social networks, employer hiring preferences, and job instability", in G. Manzo (ed.), *Analytical Sociology: Actions and Networks*. Chichster: Wiley and Sons: chapter 14.

Gallager, R. G. (2013) *Stochastic Processes: Theory for Applications*. Cambridge: Cambridge University Press.

Sawyer, R. K., 2005. *Social Emergence: Societies as Complex Systems*. Cambridge: Cambridge University Press.

Gelman, A., Hill, J. (2007). *Data Analysis Using Regression and Multilevel/ Hierarchical Models*. Cambridge: Cambridge University Press.

Girard, A. (1987). Jean Stoetzel (1910–1987). *Revue française de sociologie*, 28, 2, pp. 201–211.

Glass, D. V. (ed.) (1954) Social mobility in Britain. London: Routledge

Goldthorpe, J. (2000) *On Sociology: Numbers, Narratives, and the Integration of Research and Theory*. Oxford: Oxford University Press.

Goodman, L. (2007) "Statistical magic and/or statistical serendipity: An age of progress in the analysis of categorical data", *Annual Review of Sociology*, 33: 1–19.

Granovetter, M. (1978) "Threshold models of collective behavior", *American Journal of Sociology*, 83(6): 1420–1443.

Granovetter, M., Soong, R. (1983). "Threshold models of diffusion and collective behavior", *Journal of Mathematical Sociology*, 9: 165–179.

Hanushek, E. A., Wößmann, L. (2006) "Does educational tracking affect performance and inequality? Difference-in-difference evidence across countries", *The Economic Journal*, 116(March): C63–C76.

Hauser, R. M., Warren, J. R. (1997), "Socioeconomic Indexes for Occupations: A Review, Update, and Critique", *Sociological Methodology*, 27: 177–298.

Héran, F. (2006). Social Mobility: Then and Now. In G. Caselli, J. Vallin, G. Wunsch (eds), *Demography: Analysis And Synthesis; A Treatise in Population Studies*, San Diego, Academic Press, vol. 3, 477–493, ch. 92.

Jencks, C., Smith, M., Acland, H., Bane, M. J., Cohen, D., Gintis, H., Heyns, B., Michelson, S. (1972) *Inequality. A Reassessment of The Effect Of Family And Schooling in America*. New York: Basic Books.

Kelley, S. M. C. and Kelley, C. G. E., (2009) "Subjective social mobility: Data from 30 Nations (November 23, 2008)", In Max Haller, Roger Jowell and Tom Smith (eds), *Charting the Globe: The International Social Survey Program 1984–2009*. London: Routledge, Chapter 6.

Khuri, A. I., Mukhopadhyay, S. (2010) "Response surface methodology", *WIREs Computational Statistics*, 2: 128–149

Laumann, E. O. (1965) "Subjective social distance and urban occupational stratification", *American Journal of Sociology*, 71(1): 26–36.

Machamer, P.K., Darden, L., and Craver, C.F. (2000) "Thinking about mechanisms", *Philosophy of Science*, 67(1): 1–25.

McCall, L., Manza, J. (2011) "Class differences in social and political attitudes in

America", in G.C. Edwards III, L. Jacobs, R. Shapiro (eds), *Oxford Handbook of American Public Opinion and the Media*. New York: Oxford University Press: 552–570.

Manzo, G. (2012) Reason-based explanations and analytical sociology: a rejoinder to Boudon. *European Journal of Social Sciences*, 50(2), 35–65.

—— (2013a) Educational choices and social interactions: a formal model and a computational test. *Comparative Social Research*, 30, 47–100.

—— (2013b) "Is rational choice theory *still* a rational choice of theory?" *Social Science Information*, 52(3): 361–382.

—— (2014) "Data, generative models, and mechanisms: *More* on the principles of analytical sociology", In G. Manzo (ed.), *Analytical Sociology: Actions and Networks*. Chichster: Wiley and Sons: chapter 1.

Merton R, K. (1967) "On sociological theories of the middle-range", in R. Merton, *On Theoretical Sociology: Five Essays Old and New*. New York: Free Press: chapter 2.

—— (1968) "On the history and systematics of sociological theory", in R. K. Merton, *Social Theory and Social Structure*. New York: The Free Press: ch. 1, 1–37.

—— (1987) "Three fragments from a sociologist's notebooks: Establishing the phenomenon, specified ignorance, and strategic research materials", *Annual Review of Sociology*, 13: 1–29.

Panik, M. J. (2014) *Growth Curve Modeling: Theory and Applications*, Wiley and Sons.

Pettigrew, T. F. (2015) "Samuel Stouffer and relative deprivation", *Social Psychology Quarterly*, 78(1): 7–24.

Raftery, A.E. (2001) "Statistics in sociology, 1950–2000: A selective review", *Sociological Methodology*, 31: 1–45.

Schelling, T. C. (1971) "Dynamic models of segregation", *Journal of Mathematical Sociology*, 1: 143–186.

Smits, J., Wout, U., Lammers, . (1998) "Educational homogamy in 65 countries: An explanation of diferences in openness using country-level explanatory variables", *American Sociological Review*, 63(2): 264–285.

Smits, J. (2003) "Social closure among the higher educated: Trends in educational homogamy in 55 countries", *Social Science Research,* 32(2): 251–277.

Sorensen A. (1998) "Theoretical mechanisms and the empirical study of social processes", in P. Hedström, R. Swedberg (eds), *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.

Tulving, E. (2007) "Are there 256 different kinds of memory?" in James S. Nairne (ed.), (2007) *The Foundations of Remembering: Essays in honor of Henry L. Roediger, III*. New York, NY, US: Psychology Press: ch. 3, pp. 39–52.

Turner, F. C. (ed.) (1992) *Social Mobility and Political Attitudes. Comparatives perspectives*. New Brunswick/London: Transaction Publishers

Wintrobe, R. (2006) *Rational Extremism: The Political Economy of Radicalism*. Cambridge: Cambridge University Press.

PART ONE

History of Social Thought

1

SOCIODICY

Some Comments on Mohamed Cherkaoui's Durkheimian Approach to Social Emergence

Salvador Giner
University of Barcelona

The effort to establish a firm relationship between sociology and a moral philosophy satisfying the demands of modern science found a convincing answer in the work of Émile Durkheim. He was unambiguous in declaring that moral facts, *faits moraux*, are socially produced and determined phenomena. Yet, in spite of such a categorical assertion, Durkheim's project also entailed a philosophical interpretation of ethics which excluded reductionism, on the one hand, and relativism, on the other. Convinced that the *détermination sociale du fait moral*, as such, implied an ethical relativism which was repugnant to his deepest convictions, Durkheim faced a serious difficulty. He found a solution, however: he embraced emergentism.

All versions of emergentism assume that one or several of the properties of one level of reality are absent from the properties perceived among their components, as found at a lower level. As Mohamed Cherkaoui observes, emergence, in Durkheim, took two essentially different forms. On the one hand, it implied the crystallization of social phenomena into institutions, such as religion, law, morality, or the

financial system. On the other, it appeared in those non-crystallized forms that Durkheim called "social currents". Currents are shared attitudes leading to certain behavioural patterns, characteristic of certain individuals. Thus, the rates of birth, marriage, suicide, delinquency, which statistics and other data reveal for a given social class, nation or community, are also emergent qualities (Cherkaoui, 2008: 27–28). By assuming that a human society possesses traits such as these emergent characteristics, either institutionalized or behavioural, and that they cannot be found in the individuals that form it, Durkheim declared that society is a *sui generis* reality.

The Durkheimian solution against both moral relativism and the absolute "sociologization of morality"—of which he is wrongly and frequently accused—is grounded on two different aspects of his emergentism: his theory of the historical evolution of humanity as a moral evolution; and his conception of society as a all-encompassing category. Let us first look at this last element of what I have called Durkheim's solution. According to Durkheim, society is essentially a supraindividual entity, higher also to all the groups that make it up. It possesses the emergent quality of the widest generality, absent in all other social groupings, institutions and individual members. This new quality is that of morality. In turn, morality is inextricably linked to the sacred. Since the sacred is that which must not be violated, that must be obeyed, and the guide to all conduct, morality too must be understood as *sui generis*. It participates in the same quality that characterizes society as a whole. "Society" and "morality" are the two facets of the same phenomenon, its structural and its normative dimensions.

It is in this context that the relationship between the social dimension of morality and ethics, or moral philosophy, becomes obvious. Durkheim himself acknowledges it when he distinguishes between his own position and that of Kant. While for the latter no morality is possible without God, for Durkheim no morality is possible without society. "Kant postulates God …,—he says—we postulate society". And he concludes, just as laconically, "We must choose between God and society".

Yet, do we? Even in strictly Durkheimian terms, this question must be automatically asked, since for Durkheim God is an abstract mode, no matter how powerful, to convey the notion, often more nebulous and

hard to understand, of society. If God is society, as Durkheim affirmed, why then must we choose? It looks as though August Comte's sociolatry, his cult and *religion de l'humanité*, could have found an attenuated and more comprehensible expression in Durkheim's social philosophy. An expression, that is, more in tune with the then dawning era of the 20th century and obviously free from Comte's essentially romantic dreams, which he nevertheless couched in a positivistic guise.

Differences apart, both thinkers shared something fundamental. Both Comte and Durkheim avoided sociological relativism by asserting the moral sovereignty of society as a new categorical imperative, sociologically grounded and philosophically acceptable, since man is essentially a social animal. It was thus that one of the most secular conceptions of the moral universe ever developed, that of Émile Durkheim, ended up logically in a divinization of that which is essentially worldly, even if the conclusion was reached by the most quintessentially *laïciste* thinker of his times.

It must be assumed that Comte's *religion de l'humanité*—a phenomenon described by him in strictly emergentist terms—lies behind Durkheim's manifestly militant, secular and republican secularism. This impression is reinforced by Durkheim's efforts to avoid the positivist fragmentation of morality under the dangers posed to it by a narrow empiricism, through his evolutionism. The long march of humankind from primeval primitivism and tribalism towards the vast complexity, functional integration, intense change and acute functional complementarity of tasks of civilized society must also be understood as a natural history of mores and values. In that history Durkheim saw—though more cautiously than earlier theorists—progress at work. A painful and slow liberation from fetishism and animism, a slow substitution of penal law by civil arrangements and agreements, a development of individualism and rationality, and the growth of functional solidarity (if not always fraternity) between humans of very diverse occupation, class, and country.

This extremely well-known theory may be summed up, in Durkheim's vocabulary, as a permanent increment, through human history, of what he termed the *moral density* of society. Population density combined with a growing social division of tasks were the sources

for the historical development of moral density and complexity. It was thus that the moral cement of society in modern times came to depend much more on complementarity and mutual functional dependence between different people and institutions—what he called organic solidarity—than on any feelings of love or benevolence towards our fellow humans we may have, as had already been suggested by Adam Smith or, just before Durkheim himself, by Ferdinand Tönnies. For the latter, unambiguously, the community was ultimately the sole source of human morality. Durkheim, on the other had, would argue that primordial—in his language, mechanical—solidarity, innate in primitive societies, could not possibly exhaust all the sources of morality. His conceptions about the morality of guilds, but notably his theory of the modern *morale professionnelle* show how he made an effort to identify social sources of morality other than primitive or "elementary" ones, producing the passage from the "is" to the "ought", among humans, that Humean gap that can never be solved in strictly epistemological terms, but that can be answered sociologically (Giner, 1991).

Durkheim's organic solidarity goes beyond primordial feelings of fraternity, for it substitutes the virtue of altruism, as imagined by Comte, who coined the word. (It is today in constant use by all and sundry. It constitutes a notorious case of linguistic secularization: as "charity" becomes obsolete outside religious contexts, and its republican equivalent, "fraternity", falls out of fashion in today's more cynical times; while "altruism" becomes suspicious, "solidarity" still enjoys a certain popular favour.) The world historical progress of organic solidarity, by itself, would lead to a moral utopia in Durkheim's scheme of things, were it not for his painful awareness of a parallel progress of the contrary forces of anomie, disorder and amorality produced by another set of causes. With this important qualification, it is obvious that Durkheim's sociology entails also an evolutionary ethical theory, as had earlier been that of Comte, Marx and Spencer. It was the moral theory of a "convinced generativist", as Cherkaoui describes his methodological approach (Cherkaoui, 2005: 35–36). And, I would add, also his ontological beliefs about the human race.

Durkheim's cautious progressivism tempers any suspicions of utopianism that may be detected in his thought. He was too painfully

conscious of the "general state of legal and moral anomie in which economic—and political—life found itself" during his own time for utopianism to creep into that progressivism, or into the democratic socialism and reformism to which his own sociological convictions were subtly attached. Durkheim was prudent. He was firmly convinced that the path to a higher civilization included defeats and serious historical setbacks, and that the control of the "lower forces", *les forces plus basses*, of social life—such as the "law of the strongest and fittest"—by their submission to higher moral laws and principles was a difficult task. (A task to which he devoted much of his sociology of education. As Cherkaoui rightly says, there are very few "sociological theories that have been so widely accepted as the Durkheimian theory of educational systems" (Cherkaoui, 2008: 115).

By embracing a form of Durkheimian emergentism while remaining interested, as much as he is in the study of the sociologically crucial paradox of the unintended consequences of human action, Mohamed Cherkaoui has made a substantial contribution to our objective knowledge of the social world (Cherkaoui, 2007). Such knowledge, if we are to judge from his work on Durkheim's postulates about people's notions of the ultimate nature of reality, should include a new form of sociodicy, that is, a convincing moral justification or condemnation of the social order, its conflicts, its evils and its victims (Bell, 1996). What theodicy—from its earliest formulation in Leibniz to this day—could not explain—and indeed, will never be able to explain, trapped as it is in its own conceptual dilemma—a sociodicy perhaps could. After all, the most outstanding social theorists, philosophers and sociologists have often put forward fairly convincing, albeit, always incomplete, theories of our evils and explanations of our sufferings. Marx's was one, so was Weber's, and so was Durkheim's interpretation of anomie. A wise combination of the achievements of what is often called "conflict sociology" on the one hand and "analytical sociology", on the other, may be the best path to follow today. Mohamed Cherkaoui's contribution, it seems to me, must be situated along this promising trail. It is, after all, the research programme of the *sociologia perennis*, always undisturbed by the facile distractions of the trivial sociological impostors of our troubled times.

REFERENCES

Bell, D. (1966) "Sociodicy: A guide to modern usage", *The American Scholar*, 35(4): 696–714.

Cherkaoui, M. (2005) *Invisible Codes: Essays on Generative Mechanisms*. Oxford: Bardwell Press.

——— (2007) *Good Intentions: Max Weber and the Paradox of Unintended Consequences*. Oxford: Bardwell Press.

——— (2008) *Durkheim and the Puzzle of Social Complexity*. Oxford: Bardwell Press.

Giner, S. (1991) "Intenciones humanas, estructuras sociales: para una lógica situacional", in M. Cruz (ed.) *Acción humana*. Barcelona: Ariel, pp. 21–126.

2

ÉMILE DURKHEIM'S DEBT TO CLAUDE BERNARD

Back to the Origins of the Sociology of Generative Mechanisms[1]

Dominique Guillo
CNRS (GEMASS UMR 8598, CNRS-Paris IV)

References to Claude Bernard are often mobilized in commentaries of Durkheim's sociology and his relationship to biology. However, it is often considered as quite secondary, compared with other sources of inspiration, such as Edmond Perrier, Henri Milne Edwards, or Herbert Spencer.[2] According to these commentaries, Durkheim merely drew on a popularized version of the "experimental method". And this general reference to a notion which was once fashionable would not play a major role in his sociology. It is a fact that Durkheim quotes Bernard only but a few times. Yet, as it was stressed by Mohamed Cherkaoui in *Invisible Codes* (2005: 43), the linkage between Durkheim's epistemology and Bernard's is very deep.[3] It is so deep and central that Durkheim turns out to be a genuine precursor of the sociology of "generative mechanisms" (Cherkaoui, 2005, Chapter 2).

This article aims to assert the thesis formulated by Mohamed Cherkaoui and at sustaining it, by demonstrating that Durkheim's epistemology and sociological theory are grounded in faithful transpositions of Bernard's viewpoints in biology in the realm of sociology, and

especially in three important points: a) the experimental reasoning, b) the interpretation of the notion of function in terms of efficient causality, c) the necessity to separate the cause of a phenomenon from its function.

As Mohamed Cherkaoui also stressed it, these three borrowings depart Durkheim from any kinds of perfunctory holism or unilateral determinism from the social realm to the individual,[4] to which the French sociologist is often associated by commentators.

1. "EXPERIMENTAL REASONING"

In his *Introduction à l'étude de la methode expérimentale* (Bernard, 2008), Bernard develops thoughts which go beyond natural sciences. His goal is to set the rules of method which are constitutive of the scientific mind in general. For him, science grounds in the first place in the idea that the natural order which is visible for man is governed by an absolute determinism. Every phenomenon has a cause. Assuming that there could be phenomena without a cause would amount to putting science in contradiction with itself. In the same fashion, no scientist can suppose that a phenomenon could undergo variations without any new conditions. The hardcore work of a scientist consists in establishing causal relationships between phenomena, i.e. laws.

For all of these reasons, the truth of scientific laws cannot only have a relative degree of probability.[5] Indeed, a causal relationship has to be but absolute and necessary. Tolerating exceptions to scientific laws would result in denying the principle of causality. Certainly, the experience—the phenomenon—has a *relative* and *contingent* dimension, insofar as it consists in a relationship between our mind, on the one hand, and things-in-themselves that are unreachable *per se*, on the other hand. However, relationships between phenomena are "necessary and absolute" (Bernard, 2008: 75). They are independent of the contingent circumstances in which they come out. According to Bernard, if there is some kind of necessity in scientific laws, it is because they consist in *ideas*, linked to each other through *logical* links. As a whole, scientific laws have two sides: one that is empirical and the other that is rational. On the one side, they are propositions that refer to states of the world.

On the other, they are ideas that are likely to be deducted one according to the other by the mere work of reason. Nevertheless, this deduction is no abstract speculation. It operates on the basis of laws formerly established by science. On this second dimension, scientific laws are necessary, as much as mathematical truths are.

The conception of the scientific method that Bernard proposes is the direct consequence of such a conception of scientific law. Scientists only rely to empirical datum or experimentation in a second move, in order to confirm or infirm the idea they formerly formulated by deduction. In other words, the reason relies on experience or experimentation to validate theoretical constructions, which, besides this, have to be sustained by the mere force of their internal logic. Thus, Bernard breaks with both some kind of blind empiricism and some sort of speculative rationalism: a law is both an empirical and a rational truth. This is why he coins the expression "experimental reasoning" to label the scientific method he advocates (Bernard, 2008, Part I, especially pp. 97–104 and p. 284).[6]

Just like Bernard, Durkheim considers that the idea of science implies the idea of an absolute determinism in the unfolding of natural phenomena. For Durkheim, this principle is true for social sciences as much as for other sciences. Considering that a social fact could exist without cause equals denying the very possibility of a social science. Moreover, just like Bernard, Durkheim deems that, to demonstrate the existence of a causal link between two—social (for Durkheim)—phenomena, one has to lead in the first place an autonomous theoretical thinking which spares this causal relationship its character of necessity, and then compare the given explanation with empirical data, so as to validate it.

This is very explicitly the approach advocated by Durkheim in *Rules of sociological Method* (1895), for instance, in a short development on suicide. To him, the determination of the cause of a social fact such as suicide has to start up with the search for "concomitant" empirical "variations", because when a general fact is the cause of another general fact, their concrete manifestations will necessarily vary in the same direction, and with the same intensity. To draw on an example proposed by Durkheim, if education is the cause of suicide, the rate of one and

that of the other would both vary in a concomitant fashion. However, as Durkheim adds, these empirical co-variations—repeatedly observed concomitant variations or statistics—do not suffice to establish a causal link between two social facts. Indeed, they do not allow establishing but mere correlations, whereas a correlation does not disambiguate which variable is the cause, and which one is the effect. Moreover, two social facts can vary simultaneously, not because one is the cause of the other, but because both are the effects of a same cause, i.e. a third variable. To put it in a nutshell, if the co-variation of two phenomena is a necessary condition to establish the existence of a causal link, it is not a sufficient condition. Then, how to establish a causal link? How can one know if two correlated phenomenon are brought together by such a link?

The answer given by Durkheim to the latter question is identical to Bernard's one: it is theoretical deduction that allows us to settle things. Let's take the example of suicide once again. Durkheim says we note a statistical "concomitant variation" between the level of education and the suicide rate. Yet, the first item cannot be a cause of the second. Why? Because "it is impossible to understand how education can lead to suicide; such an explanation contradicts the laws of psychology." (Durkheim, 1988: 224). For Durkheim, only deduction allows establishing a causal relationship between two phenomena. For this reason, the method that authorizes the administration of the proof is the following:

> First we shall discover, with the help of deduction, how one of the two terms was capable of producing the other; then we shall attempt to verify the result of this induction with the aid of experiments, i.e. by making fresh comparisons. If the deduction proves possible and the verification is successful, we can therefore regard the proof as having been demonstrated (Durkheim, 1988: 152).

As we can see, the method proposed here by Durkheim matches very precisely with Bernard's "experimental reasoning". For one as for the other, scientists have to begin with the search for a rationally relevant interpretation of the link between the phenomenon under study and their alleged cause—and it is this interpretation that will grant a

character of logical necessity to the relationship—and then cross-examine it with empirical data. Moreover, both Durkheim and Bernard underline the fact that it is not necessary to multiply observations to produce evidence. A well-chosen experimentation with a proper interpretation is enough for the empirical validation of a hypothesis (Durkheim, 1908).

Besides, the answer Durkheim gives about the link between suicide and education shows unambiguously that his thinking on this question cannot be reduced to a perfunctory holism. Indeed, Durkheim goes as far as asserting that one has to refer to the laws of psychology to discriminate between correlations repeated at a sociological level. Of course, this answer does not mean that Durkheim was a reductionist sociologist. Here again, there's a deep proximity with Bernard. As Mohamed Cherkaoui has stressed it, this answer means that Bernard—and therefore Durkheim—did both pioneer the sociology of "generative mechanisms" (Cherkaoui, 2005: 93; 2008). On the one hand, correlations, just like causality, link directly *social* facts. The laws that derive from them are thus *sociological* laws in the fullest sense of the word, just like the laws that govern the regulation of glucose within the organism. On the other hand, as we shall see through the very specific conception of efficient causality that Durkheim also takes up from Bernard, actions and retroactions from a level on another (whether sociological, psychological, biological, and even physical and chemical) are extremely numerous and intertwined. Consequently, in order to explain how a cause produces an effect, one has to go back to the study of extremely complex causal mechanisms. To sum it up, ontologically, social phenomena are reduced to "physical and chemical" causes. But, from an epistemological and a methodological perspective, and because of this complexity, it is necessary to consider that they enjoy a form of autonomy *vis-à-vis* proper phenomena of inferior levels of complexity, more or less intensely according to the cases considered.

This epistemological model accounts for the reason why, according to Durkheim, the complete explanation of social phenomena goes beyond the establishment of a direct causal link between phenomenon A and phenomenon B, without any references to lower phenomenal levels, as asserted by holistic commentaries of the French sociologist. This implies that one has to search for the causal chains—the mechanisms—through

which A produces B, and this imposes to go deep in the intertwinements between micro- and macro-levels. In this sense, this implies searching for social phenomena's "generative mechanisms" (e.g.: Cherkaoui, 2008).

Durkheim's interest for such a research field that mixes variables of different levels of complexity is so clear that he went as far as granting it a name, in a too often forgotten excerpt of *The Division of Labor in Society*: "socio-psychology".[7] Such an interest does not contradict Durkheim's thesis on the epistemological autonomy of social matters, which is often underlined. Indeed, one can consider that social sciences are epistemologically independent from psychology and biology, and yet advocates that the problem of articulating accumulated knowledge in different disciplines is a problem of a very specific kind, which does not have an immediate answer, such as "social aspects determine the individual integrally" or "biologic aspects determine the individual integrally". In this regard, unlike some of his modern interpreters, Durkheim did not mingle the thesis of sociology's epistemic independence on the one hand, and the thesis of individuals' social determination on the other. Here again, his sociological thinking echoes Bernard's physiology. And through it, it gets even closer to the generative sociology.

2. THE INTERPRETATION OF FUNCTION IN TERMS OF EFFICIENT CAUSALITY.

As we know, in order to build up his theory of society, Durkheim grounds his first essays on numerous organist analogies. As we will see, the model of organism he relies upon is very close to the one proposed by Claude Bernard. In the first place, what Durkheim draws on from this model is the extremely original conception of efficient causality which constitutes the epistemological foundation of Bernard's physiology.

In the pages he dedicates to the general theory of the organism, Bernard says he wants to operate a synthesis between the two major paradigms which oppose in physiology. The first one is often supported by German scientists. According to them, the organism is a synthetic totality controlled by a "vital force" irreducible to an analytical decomposition. Consequently, vital phenomena—for instance, the liver's functioning and its *raison d'être*—must be fully explained by the specific

role assigned by this force to the organism's general harmony. They thus cannot be explained by a principle of efficient causality as it is used in physics. They have to be explained by a final cause (e.g. Bernard, 2008: 122-126). In the second paradigm, on the contrary, the organism is considered as a physical aggregate as any other. It only proves to be of a larger complexity than inert aggregates. To account for a physiological fact, scientists have to push on the process of analytical decomposition as far as possible. In other words, they have to link directly the fact under study to the physic-chemical laws which govern the ultimate components of the organism (Bernard, 2008, Part II, Chapter 2).

For Bernard, both paradigms encapsulate part of the truth; but they also contain a symmetric mistake. On the one hand, as Bernard considers in reaction to the supporters of the "synthetic model", science has to give up on any references to the "vital force" and, more generally, to any forms of final cause. Such blurry entities stand in contradiction with the scientific spirit. They can explain everything and, thus, they shed no light on anything. Nothing can withdraw from the principle of efficient causality. Thus, just like the supporters of the "analytic model", Bernard asserts that physiological facts ultimately ground in the laws of physics and chemistry (Bernard, 2008: 124-26). However, to him, defenders of this model are wrong when they do not take into account the specificities of the organism in regards with other physical aggregates. Not only does the "determinism" to which physiological phenomena are submitted is very complex but it is also "harmonically hierarchized" into successive levels of integration (Bernard, 2008: 165). Complex physiological phenomena are constituted by a series of simpler physiological phenomena, ordered and coordinated one to the other in such a way so as to produce a particular effect. Consequently, for Bernard, physiologists should always recompose the "physiological *synthesis*" (Bernard, 2008: 165) of the phenomena under study, after decomposing it through the analysis, in order to unravel the concerted action of all of the parts which have been previously isolated.

In sum, for Bernard, determining the activity of an organ—i.e. its function—(the very object of physiology) lies upon a very original epistemology. This latter narrowly links the principle of efficient causality and the principle of the organism's intern harmony. For Bernard, on an

ontological level, physiological phenomena are nothing but the product of physical and chemical actions and reactions. Yet, epistemologically speaking, one has to mobilize an analytical approach *and* a synthetic one in order to understand them. On the one hand, what is needed is an approach that aims at isolating some mechanisms—possibly on a micro (i.e. chemical) level—implied in a physiological phenomenon—i.e. a macro-phenomenon—that scientists have to account for. On the other hand, one has to demonstrate how the isolated mechanisms contribute to the production of the phenomenon through a series of mutual actions and reactions. This is the meaning of the "physiological synthesis" to which scientists have to proceed. Even though it operates at different levels of complexity and integration, this synthesis is always governed by the principle of efficient causality. But it deals with phenomena of a specific kind: they interact one with the others and retroact *on themselves*, through the chain of effects that they produce.

In a surprising excerpt, Bernard develops this idea by using the same kind of operation used by organicist sociologists, but the way around: the analogy between organism and society. One can find in his writings the very same words that will be used by Durkheim later on to characterize modern societies. As Bernard puts it, there is an "organic or social solidarity" (Bernard, 2008: 166) between physiological phenomenon. Indeed, these phenomena are interconnected by chains of efficient causality relationships, just like physical and chemical phenomena. Yet, in the case of the organism, these chains are self-sealing, circular. In other words, a physiological phenomenon produces effects on itself through the chain of its consequences. The "vital *circulus*" produces a balancing of the whole, at different levels of integration, through reciprocal causal actions of the elements of a *system*. In this model, the function does not appear as a final cause. It is henceforth a mere portion of a long and ramified chain of efficient causes which presents the particularity of being closed on itself. In sum, the study of a physiological phenomenon has two sides which have to be carefully differentiated.[8] Consider a physiological phenomenon (X)—i.e. a macro-organic phenomenon like the functioning of the liver. This phenomenon (X) is the product of a physiological cause (Y). This latter produces its effect (X) through extremely complex mechanisms which imply most of the time

several levels of integration of the organism. On the other side, this phenomenon (X) produces a chain of effects (Z) which can little by little have consequences on phenomenon (X) itself. This is what is called the function of phenomenon X—for instance the liver.[9]

This is exactly the meaning that Durkheim grants to efficient causality and the notion of function in sociology in *The Division of Labour in Society* (1893) and *the Rules of Sociological Method* (1895). In these two works, the French sociologist stresses that the principle of efficient causality as it is used in physics is not enough to account for social facts. As for living creatures, societies are governed by a principle of internal equilibration. Durkheim underlines that, consequently, the sociological explanation has to underscore the *function* of phenomena, in addition to the search for their *cause*. But the French sociologist cares to point out, like Bernard, that the notion of function has to ground in a principle of circular efficient causality which disqualifies any finalistic interpretations of the role of social facts—in particular institutions and norms.

> If we have opted for the term [of function], it is because any other word would be inaccurate or equivocal. We cannot use the words 'goal' or 'object' and talk about the end of the division of labour, because it would imply that the division of labour exists *in order to accomplish the results* that we will determine. The words of 'result' or 'effect' would not suit any better, because it does not echo with any idea of correspondence. On the contrary, the words 'role' or 'function' have the great advantage to imply this idea, without prejudicing the question of knowing how this correspondence has emerged, i.e. if it results from an intentional and preconceived adaptation or an afterwards adjusting. (Durkheim, 2007: 11–12)

For Durkheim, the search for the function of a phenomenon does proceed from the principle of efficient causality. But it has its own specificities, just like the search for the function of a biological phenomenon. In order to determine it, one should not focus on the elementary units—individuals and their intentions, in the case of sociology—which compose the whole. It would amount to committing the reductionist

mistake which consists in simplifying in an exaggerated manner the mechanisms which links together the micro-level of individual decisions to the macro-level of social facts. But one should not focus on the totality either, and deduce the function of a social fact on the basis of considerations on its assumed usefulness for the survival of the system. In other words, for Durkheim as for Bernard, the notion of function has to been purified from any reductionist or finalistic connotations. It means a correspondence or a mutual support between social facts within a given society, without prejudice to the causes which could have produces these social facts. Questioning the function of a social phenomenon then consists in showing to which need it "matches correspondingly" (Durkheim, 2007:11), and not to which *intention* it responds. To put it in a nutshell, for both Durkheim and Bernard, the functional explanation relies on the principle of "solidarity" between the cause and the effect:

> The effect cannot exist without its cause, but this cause needs its effect. We would not wish to raise questions of general philosophy which would be inappropriate here. However, we note that, if more closely studied, this reciprocity of cause and effect could provide a means of reconciling scientific mechanism with the teleology implied by the existence and, above all, the persistence of life (Durkheim, 1988: 188–189—nbp).

3. THE NECESSARY DISTINCTION BETWEEN CAUSES AND FUNCTIONS

Furthermore, in order to erase any forms of finalism in social sciences, Durkheim insists on what he considers to be a major necessity: to distinguish sharply between the cause of a phenomenon on the one hand, and its function on the other. Through this distinction, he targets thinkers who make out of the—real or assumed—utility of a social fact, the cause of its apparition. This critic especially targets the utilitarian school. According to Durkheim, the existence and the persistence of social phenomena such as the division of labour have often been accounted for by representatives of this school as the effects of enlightened voluntary

actions. These enlightened wills would have perceived all the benefit that can be drew on from such social phenomena, like institutions. These latters would thus have been created with regards to these benefits. Yet, as Durkheim stresses in a famous text, "this method confuses two very different questions. To demonstrate the utility of a fact does not explain its origins, nor how it is what it is. The uses which it serves presume the specific properties characteristic of it, but do not create it" (Durkheim, 1982: 119–120).

Drawing on arguments that can be found in Bernard's works, Durkheim adds that this type of explanation is invalidated by a phenomenon that one can observe in both biology and sociology:

> A fact can exist without serving any purpose, either because it has never been used to further any vital goal or because, having once been of use, it has lost all utility but continues to exist merely through force of custom. There are even more instances of such survivals in society than in the human organism. There are even cases where a practice or a social institution changes its functions without for this reason changing its nature (Durkheim, 1982: 120).[10]

Through these remarks, Durkheim comes back to the fullest meaning of the Darwinian theory of natural selection; i.e. the necessity to distinguish clearly between the causes of the apparition of a phenomenon, on the one hand, and that of its persistence through the times, on the other:[11] "when one undertakes to explain a social phenomenon, the efficient cause which produces it and the function it fulfils must be investigated separately" (Durkheim, 1982: 123).

Thus, the explanation of a social phenomenon must split into two different operations which are both combined but distinct: a search for its origin and a search on the way it succeeds in persisting as time goes by, through the ramified chain of benefic effects that it holds retroactively on its own existence. Thus, explaining a phenomenon by its function consists in explaining it by its consequences. On this topic, the filiation with Darwin is explicitly acknowledged by Durkheim himself:

> In societies as in the individual, changes occur, and these have causes, but not endings, something analogous to Darwin's individual variations. Some of them can happen to be useful; but this usefulness is not preconceived and was not a determining cause (Durkheim, 1887).

In other words, social phenomena like biological ones occur by the virtue of their causes without any relation with their usefulness. When they are beneficial for the organism, through the complex chain of their consequences, they pursue their existence. No individual or collective intention can explain directly this persistence. Through this distinction between causes and functions, Durkheim paves the way to what will be called later on in the biology of evolution *teleonomic*-type models (Mayr, 1982), i.e. explicative models which describe phenomena of internal or external harmony—adaptation to an environment—as composition effects of blind causal processes without proper finality.

In sum, one can recover in a very precise way Bernard's different topics in the sociology developed by Durkheim in his first writings about causality in general and function in particular; namely: the concrete and definite adhesion to the principle of efficient causality, a reductionist ontology nonetheless associated with an antireductionist epistemology. However, as for Bernard, this antireductionism—vigorously supported by Durkheim –, cannot be considered as a kind of holism. As illustrated by the example of suicide, for Durkheim, sociologists have to explain how a social fact causes another social fact by relying on a toing-and-froing between analysis and synthesis, between actions and reactions of the different levels of integration the ones on the others. In other words, as Mohamed Cherkaoui stresses it with strength (Cherkaoui, 2005; 2008), the complex and multidimensional chain of *mechanisms* which contribute to producing a macro-social phenomenon—as much for the causes that directly generate it and for the functions it fulfils—holds a central place in sociological explanation for Durkheim.

NOTES

1. This research has been realized with the support of the ANR (French National Research Agency) LICORNES and ASCE progammes.
2. Berthelot's (1988) very instructive analysis stands as an exception to this.
3. "Contrary to what has been argued by several commentators on the intellectual links between the biologist and the sociologist, Durkheim as not in all likelihood simply retained the importance of the hypothesis, the role of determinism, applied rationalism—fairly ordinary ideas, it must be said—Bernard taught him in particular that demonstration of laws that Comte saw at the ultimate aim of science, was only a stage that research must go beyond, that it is necessary to account for, that is to say to go further than the empirical correlation between two phenomena" (Cherkaoui, 2005: 43).
4. On this point, see also Berthelot (1988).
5. On all of these topics, see Bernard (2008, Part I).
6. On this topic, see also Berthelot (1988).
7. "In that there is a vast region of consciousness the genesis of which is unintelligible by the mere psychophysiology, it cannot be concluded that it has formed on its own and, following this, that it is refractory to scientific analysis, but only that it falls under another positive science that we could call socio-psychology. The phenomena which constitutes this latter discipline have a mixed nature: they have the same essential characteristics as other psychic facts, but they proceed from social causes" (Durkheim, 1893: 341).
8. The necessity and the consequences of this distinction are developed in the upcoming section.
9. On all of these topics, see Bernard (2008, Part II, Chapter 2).
10. It is worthy to note that Durkheim's stance implies that the nature of function should not be deduced anymore from the nature or the morphology of an organ. On this point, once again, it can be evaluated how important have been for Durkheim physiologists like Georges Cuvier or Claude Bernard, who have contributed to delocalize the function and set physiology free from the burden of anatomy. See for instance Guillo (2003).
11. On this point in Darwin's theory, see in particular Mayr (2001).

REFERENCES

Bernard, Claude (2008 [1865]) *Introduction à l'étude de la médecine expérimentale*. Paris: Flammarion.

Berthelot, Jean-Michel (1988) "*Les règles de la méthode sociologique* ou l'instauration du raisonnement expérimental en sociologie", introduction to Emile Durkheim, *Les règles de la méthode sociologique*. Paris: Flammarion, pp. 7–67.

Boudon, Raymond (1999) "*Les formes élémentaires de la vie religieuses:* une théorie toujours vivante", *L'Année sociologique*, 49(1): 149–198.

Cherkaoui, Mohamed (2005) *Invisible Codes*. Oxford: Bardwell Press.

——— (2008) *Durkheim and the Puzzle of Social Complexity*. Oxford: Bardwell Press.
Durkheim, Emile (1887) "La science positive de la morale en Allemagne, extrait de la *Revue philosophique*, 24", in Emile Durkheim (1975) *Textes*, vol. 1. Paris: Editions de Minuit, p. 289.
——— (1908) "[Debate on explanation in History and Sociology], extrait du *Bulletin de la société française de philosophie*, 8", in Emile Durkheim (1975) *Textes*, vol. 1. Paris: Editions de Minuit, p. 203.
——— (1988 [1895]) *Les règles de la méthode sociologique*. Paris: Garnier Flammarion.
——— (2007 [1893]) *De la division du travail social*. Paris: PUF.
Guillo Dominique (2003) *Les figures de l'organisation. Sciences de la vie et sciences sociales au dix-neuvième siècle*. Paris: PUF.
Mayr, Ernst (1982) *The Growth of Biological Thought*. Harvard: Belknap Press.
——— (2001) *What Evolution Is*. London: Penguin.

3

ENGINEERING THE WELFARE STATE

Gunnar Myrdal, Social Science and the Politics of Unintended Consequences

E. Stina Lyon
London South Bank University

> "It is really true what philosophy tells us, that life must be understood backwards. But with this, one forgets the second proposition, that it must be lived forwards."
> —Soren Kirkegaard, (*Soren Kirkegaards Skrifter*, volume 18, p. 306, 1843)

INTRODUCTION

Discussions about unintended and unanticipated consequences in sociological research practice have largely fallen by the wayside in the history of sociological research since Merton wrote his famous essay on the topic (Merton, 1936). Their theoretical definition and understanding by the sociological classics have been brought to the fore by Cherkaoui in his recent works on Weber and Durkheim (Cherkaoui, 2007 and 2008). But their practical implication for what the Swedish social scientist Gunnar Myrdal, in a critical comment on Weber's grandiose "speculations", called a "working social scientist", has

gained less attention (Andersson, 1999). This is particularly so when it comes to historically understanding the research processes that underpinned welfare state developments and that were aimed democratically to encompass human behaviour and values in all their contradictory and unpredictable complexity. For a "working social scientist" like Myrdal, a member of major policy preparatory commissions in Sweden during the 1930s, the relationship between collective facts and individual values, scientific and normative reasoning, hypothetical predictions about outcomes, and opposing ideologies about the nature of means and ends were so closely intertwined that issues of complexity could not be put aside as theoretical anomalies to be dealt with later. Instead such complexities had to be exposed as value laden dilemmas from the start of whatever evidence collection was seen as necessary as value laden dilemmas to ensure the economic efficacy and moral and political acceptability of social reform. In his view, the best a social scientist could do was to try to describe, explain and predict, as validly and dispassionately as possible, the complex social circumstances under scrutiny whilst being honest about his, or her, own value position so as not to mislead with hidden, ever false, ideological certainty. Myrdal's own particular social democratic "point of view", to choose his own term, can be seen both as a product of and a contributing factor to Swedish welfare state developments and their varying status over the last century as both a model of democratic progressive and egalitarian reforms, and an example of the inherent moral dangers of state planning for social engineering purposes (Carlson, 2005; Barber, 2008). The methodological dilemmas outlined in his work, dilemmas arising less from any particular "partisanship" than from the nature of social science itself, remain as pertinent as ever, as do questions regarding the nature of welfare state planning and its often contradictory consequences.

Myrdal, an economist by training, but also well educated in jurisprudence, sociology, social psychology and education, was an active and significant participant in the development of Swedish social politics during the 1930s, a period of modernisation, rationalisation and social engineering led by the Social Democratic Party.[1] At the time he was already internationally well- known for his economic writings on business cycles and values in classical economic theory (Myrdal, 1953a).

This visit coincided with the Wall Street crash, an event to kick start a decade of devastating depression. After a year on Rockefeller grants to the U.S. with his wife and intellectual partner and frequent co-author Alva, he was also a newly converted "sociologist", partly due to his intellectual friendship with W. I. Thomas, whose work on the significance of "social definitions" in explaining action he greatly admired (Myrdal, 1944: lxxxi).[2] Myrdal contributed to the reform process through his membership in several royal and government commissions, a uniquely Swedish tradition of bringing social research to social reform through the engagement of subject expertise, interest group and political party representation. Government research commissions (SOUs) were the forum for presenting and rationally critiquing social research evidence. As Wisselgren has argued, this was where early social scientific culture and politics met and spheres of action for expertise and expert knowledge were navigated, defined and regulated and knowledge-based social reforms instigated. As policy preparatory organs, the commission reports were also aimed to facilitate consensus in parliamentary decision making. During the early period of Social Democratic rule and its ambitious reform programme, their number and individual length increased (Wisselgren, 2008). The ensuing Swedish welfare state reforms in the interwar period of modernisation and rationalisation can be seen as a more democratic model of social engineering than welfare models developed at the same time in the totalitarian states of Germany and the Soviet Union. It was a unique period in the creation of transformative social and economic policy and unusual when it came to the close relationship between intellectuals, social scientists, and political decision making (Eyerman, 1985; Olsson, 1990; Berggren and Tradgardh, 2009).

In scrutinising a few of the commission reports to which Myrdal made a direct contribution with respect to the problem of social complexity and unintended consequences, the aim is to focus on what Kaplan referred to as the "logic in use" of his methodology, rather than attempting to present a "reconstructed logic" of his presentations in terms of policy content or outcomes (Kaplan, 1964). For Myrdal to try to scientifically anticipate possible consequences of social and economic reforms, and research techniques to evaluate them, including of the complexity caused by changing, and changeable, vagaries of political

priorities, public opinion, and shifting external social and economic conjunctures, became a core researcher task. It was a task, however, that he experienced as politicised as the goals and values of the social reform programmes to be developed. The goal of value free social scientific analysis became for him not only a practical *chimera*, it was also, given the misery and injustice of the human condition, ultimately an immoral one. Such a position of reflexive value self-exposure led him to write not only with great conviction but with a strong and persuasive argumentativeness that creates a way into understanding the complexities of social science as an applied *moral* science. He saw as one of his "expert" tasks as that of educating the public through writing clearly and accessibly on the problems of bringing social science information, methodological rigour and logical reasoning to everyday policy decisions, and of its role in moving from scientific "prognosis" of likely events to a political "programme" realisable in practice. He did so, sometimes within requested summary research reports, and sometimes in appendices, the most famous of which appear in his later *opus magnum* on race relations in the U.S. (Myrdal, 1944: Appendix 2; Lyon, 2004).

SOCIOLOGICAL RESEARCH AS THE ART OF THE POSSIBLE

Why is revisiting these commission reports a worthwhile task from the point of view of sociological history? The concept of an academic sociology as latterly understood in the discipline has changed and was not equally defined during the earlier half of the last century, nor equally constructed in different national contexts. The history of sociology is a great deal more fluidly interdisciplinary, and more mixed when it comes to conceptions of science and research methods than traditionally given credence by post-war discipline purists. The notion of a cumulatively developing stock of knowledge arising from rule bound critical exchanges within well defined discipline boundaries, a perspective originally formulated by Merton (1957) is a misleading one, given the way individual social scientists tend to respond to and engage with political events and circumstances, particularly in times of crisis (Wagner, 2001). As Honegger and Wobbe argue in their discussion of women

in the sociological tradition, in the institutionalisation of a discipline there is a selective aspect with certain traditions kept out of the way and functioning as a negative backing "foil" for the drawing of boundaries in the competition for space, status and attention between, and within, academic fields (Honegger and Wobbe, 1998). In this process, argues Lepenies, scientific programmes having once been rejected do not simply disappear, but pass a period in "storage" from which they again may return (Lepenies, 1988). One such programme is that of the democratic welfare state and the role of research within it with notions such as "the middle, or third, way", "social engineering", or even "state" functioning using McLelland's term, as "vehicular concepts" travelling and changing at speed to adapt to altering political circumstances (McLelland, 2004; Podgorecki, *et al*, eds, 1996).

But "real world" sociological research as an intellectual and empirical activity has remained coherent enough to enable Smelser to sum up its intellectual "field" as having three basic orientations: a scientific one emphasising causal analysis and dispassionate empirical investigation, a humanistic one emphasising preoccupation with the human condition, including social welfare, justice, equity and human meanings, and finally an "artistic" one emphasising the craft of application of sociological knowledge, as in "the art of the possible" (Smelser, 1997). Much debate has taken place over which one should have primacy in the sociological task, and over their inherently contradictory nature (Burawoy, 2005; Turner, 2007). They well summarise the task Myrdal set himself in bringing sociological as well as traditional economic evidence to bear on furthering the cause of a more fair and humane social world and provide a framework for understanding the many "dilemmas", a frequently occurring concept in both the Myrdals' writings, inherent in a broader and more complex and applied conception of the nature of the discipline.

THE COMPLEXITY OF SOCIAL LIFE AND ITS CONSEQUENCES

In revisiting the sociological classics to help understand the contemporary world and the potential for critical intervention in its development,

Cherkaoui has alerted us to the long-standing problem of social complexity and the paradoxical relationship between intentions, actions and outcomes in social affairs. In his work on the proverbially weak relationship between good intentions and their unexpected outcomes, Cherkaoui provides a systematic account of Weber's writings on religion and its relationship to the development of capitalism (Cherkaoui, 2007). At the core of Weber's work lies the proposition that the inner-worldly asceticism of protestant values created a set of aspirations and behaviour expectations that had, if not a direct causal effect, a distinct elective affinity with the economic rationality of capitalism as a driving force of social and economic change. Frugality and asceticism paradoxically led to wealth creation, the economic engine that fuelled economic growth and expansion. Ideological belief did have material consequences, but not always as intended or foreseen. On the basis of Weber's writings, Cherkaoui develops a model of good intentions which outline the complex relationship and tensions between intentions and outcomes, means and ends. This model raises issues of relevance both for the understanding of the limitations of social research as an explanatory and predictive enterprise and of the fragilities inherent in policy planning as an exercise in the application of social science to real world affairs. Cherkaoui notes the similarity of this model derived from Weber's writings with that presented by Merton in his classical essay, and parallels will be drawn below (Merton, 1936).

Cherkaoui outlines five types of sources of the emergence of unintended or unanticipated consequences. These "types" will provide a loose frame for my reading of Myrdal's commission work. Analytically distinct, they have the potential of creating combinations amongst themselves that produce mixed forms that are, as Cherkaoui notes, frighteningly complex to analyse. Loosely summarised they are the following: First, a mismatch between the means employed and the ends pursued caused by ignorance or error leading to wrong predictions. Poor or insufficient information about an "open-ended" social system and the inherent difficulties involved in social measurement and calculations may lead to reliance on less than rational procedures such as tradition or social imitation, or as Merton saw it, social knowledge is always less than perfect in that the most that can be established are

stochastic relationships that do not allow for a logical derivation of the relationship between intentional action and outcomes. Second, the structure of dependence between social actors means that actions will be directed to the past, present or future actions of others, which are all external to the action itself. In policy contexts, as Merton argued, this means that the urgency of immediate interests directs attention to short-term consequences at the exclusion of other more long-term possible results of the action. Third, despite a good match between means and ends the action spills over the boundaries of the domain in which it is aimed to apply and gives rise to unwanted and unintended effects in those other domains. Forth, what may be seen as rational in one social organisational sphere may appear as irrational in another leading to conflict over the meaning and evaluation of outcomes. Such conflicts cut across different social domains, but also across discipline boundaries for between economics and politics for example, or politics and religion. Fifth, a sudden emergence of new values, or an ideological resurgence of old ones, may occur during the processes of carrying out a decision creating conflicts with the values on which it was based. Such conflicts of values may at times seem irreconcilable and insoluble. For Merton this is the underlying aspect of what he saw as the self destruction of predictions, one if which is the now proverbial self-fulfilling prophecy. The prediction fails to materialise, not because it is false or improbable, but because once stated it becomes part of what is acted on thus leading to its never taking place. Or it may become true, but for other reasons than those intended and anticipated. The study of unintended and unanticipated effects is not enough as illusory expectations may keep real effects from view.

Amongst the paradoxes outlined above, argues Cherkaoui, lies the antagonism of values, the pluralism of which is a fact of life that cannot be ignored, and a social complexity that leaves both society and social science in an anxiously precarious situation with respect to efforts at asserting some kind of order and control. In his work on Durkheim and the puzzle of social complexity, Cherkaoui uses Durkheim's history and analysis of education systems to emphasise that this value pluralism over intentions and outcomes is not a random one (Cherkaoui, 2008). It is on the contrary a highly structured and organised system of

opposing interests in conflict not only over material resources and their distribution but also over the "categories of thought" in terms of which society, and individual motives within it, are to be understood. As educational reform movements in history have brought complex ideological and material conflicts, so have movements for social welfare reform, leaving social science researchers in methodological as well as political value dilemmas.

SOCIOLOGICAL PASSIONS AND THE "ROMANCE OF THE ENGINEER"

Enthused by their first contact with American sociology, but also shocked by the dire immediate social consequences of the Wall Street crash, the economist Myrdal came to embrace a conception of sociology as an integrative discipline capable of bringing together different theoretical and methodological perspectives in a more unitary fashion to support social reform. Society as a complex entity of tradition bound institutions and interacting individuals could not be understood, let alone changed for the better, without taking a multidisciplinary approach. In 1930, Myrdal wrote a passionate letter to his Danish legal colleague Alf Ross. In this he explained his conversion to sociology and his disappointment with economics as a tool for change. He wrote of his and Alva's plans for a new project: a socio-economic study of social politics, including its financing. The letter is worth quoting at some length in the way it throws light on the conception of sociology as an all encompassing approach to the "social" and its role in democratic politics. The project, he wrote, will,

> … be a critique of Marxist historical philosophy, not only a logical one, but I want to sort out its purely "positive" scientific part, especially the concept of "necessary development" and its association with teleology, which makes it unable to explain the reality of rapid social change for individuals and communities. But if one adds to this a few ounces of modern social psychology, including "patterns" and "cultural factors", one gets a modern sociology of a very realistic kind. I want to show how

> changes in themselves create tendencies towards integration within organisations, up to state and legal level (organs of "acclimatisation" in this ever faster changing world). I want to look at the sociological problems of organisations, leadership, opinion formation, psychological problems, interests, attitudes and different degrees of social power. All this is a *dynamic* problem, and it is the correlation between different kinds of changes that I want to catch, changes in the so called material conditions, social organisation, ideology and value attitudes (changes falling under the headings of different social sciences.) Certain generalisations will be possible, whilst retaining a certain degree of realism. Rapid change breaks down steady "folkways and mores" and people learn to think in terms of relevance (from the perspective of particular interests and attitudes), economy and *compromise*. People organise themselves to get protection from the wind and to realise important essential interests. Already in individuals doubt and compromise arise, and organisations *negotiate in changing situations*. This is how absolutist societal and psychological conceptions about what is seen as "right", that under less moveable circumstances have their base in long standing traditions, get broken down. …People "regulate" and reach compromise over what is no longer seen as "natural" and over things they see as essential for the more factual and realistic content of rights: this is the significance of democracy as a means of dissolving absolutist social phenomena. (Letter in Myrdal Archive, 23.1.1930: my translation.)

This is a far from modest proposal in its vision of sociologically capturing the political imagination of a nation in a state of social and economic flux, yet steeped in outmoded and rigid political traditions. Myrdal's first political contribution was to join his wife and become a member of the Social Democratic Party in 1932 (Jackson, 1990). In the same year he laid down his political marker in a two part paper in the Swedish modernist cultural journal *Spectrum*: "The Dilemma of Social Politics" (Myrdal, 1932). In this paper, directed at a broad audience, he emphasised the need for transformation of welfare politics to break the

moral dilemma created by the political and ideological impasse over unemployment and the failure of economic growth by itself to overcome it. Political tension in Sweden was rising, with the military called in against worker demonstrators with lethal consequences. The paper is a plea for a "middle of the road" ideology as a guide to political action to avoid the danger of social and political breakdown as witnessed elsewhere in Europe. Failure to act, he argued, carried politically dangerous consequences in disaffection and sharpened internal political conflict, an examples of which the Myrdals had experienced firsthand on visits to Germany. He asks "reasonable" people to sit around the table and come up with a compromise offering some hope for the future, especially for poorly educated young people without employment or housing prospects. A multi-disciplinary approach was called for, he argued, to include sociological assumptions about a series of issues: the negative effects of industrialisation on economic distribution; the "cultural lag" in social and political responses to rapid technological change and the need for educational reform to remedy this; the changing demands on social organisations and institutions including the family.[3] Classical liberalism, an inherently conservative ideology with respect to the sanctity of property, was unable, because of its faith in the positive effects of a free market, to resolve basic needs for redistribution when required. Classical Marxist socialism, on the other hand, had also been unsuccessful in developing an effective approach to social politics. Its prediction of first a bourgeois then a proletarian revolution, after which social politics would not be necessary, treated welfare politics as evidence of the capacity of capitalism marginally to adapt itself to changing circumstances and therefore to be ignored. Both Marxism and economic liberalism thus made strong scientific claims about a future they could not possibly know and a different kind of analysis is necessary. An ideological vacuum of political frustrations offered the possibility of compromise made possible because of common roots in Enlightenment reasoning: a socialistically softened liberalism abandoning its resistance to the regulation of private property, especially housing, for the greater collective good, and a liberally softened socialism abandoning its extreme class conflict theories in favour of accepting citizen supportive measures in the here and now.[4] Like other "third way" thinkers such as Oppenheimer and

Mannheim, and later Giddens, Myrdal assigned a special role to public intellectuals and scientifically informed social research in advancing such a reasoned compromise (Oppenheimer, 1962 [1919], Mannheim, 1972 [1929] and Giddens, 1998). The argument that the preservation of democracy depended on a pragmatic and unifying moral consensus was not a new one, nor a particularly Swedish one at the time, but the Swedish Social Democratic intellectual elite network was considerably smaller and a great deal closer to government than elsewhere at the time A political culture of democratic pragmatism and faith in the state as guarantor of individual autonomy and social justice, including for women, was historically well entrenched in Sweden making a positive reception more likely (Berggren and Tradgardh, 2009). In this "meeting in the middle", democratic discussions about the achievement of democratically established economic and welfare goals would require a more systematic research approach in fields such as education, housing, unemployment, crime and industrial economic organisation. There was, he argued, the need for a new kind of publicly accountable knowledge "expertise" to serve in the creation of a theoretical multidisciplinary synthesis, in improved empirical fact finding and in improved understanding of implementation for the sake of the preservation of democracy. His arguments were, he argued, inspired by the "romance of the engineer" applied to social rather than technical problems. One of the tasks of such a social engineer was to find a path through social complexity and the myriad of possible actions and consequences to which this complexity gave rise.[5]

PREDICTING THE FUTURE AND THE ART OF PLANNING

Unpredictable Business Cycles and Planning for their Consequences

The Social Democratic electoral growth in 1932 and the fresh economic crisis of the same year led its then leader Per Albin Hansson to redefine the party as a "people's" rather than a workers' party and to present a "crisis" programme of emergency labour and unemployment relief. It also led to the Myrdal's direct association with the leadership of the

party. Myrdal having previously written extensively on "the theory of economic fluctuations," one of his first tasks was as a member of a commission on the cyclical nature of unemployment.[6] A member of the Stockholm School of economists at the time, he was in 1933 asked to produce a memorandum on the case for an expansionist fiscal policy and on the dangers of maintaining a balanced budget in periods of high unemployment and a reduced tax base. The writing of this report forced a direct confrontation with the complexities of unintended and unanticipated effects. It also reflects the Stockholm economists' search for greater scientific rationality in the methodology of planning processes themselves rather than a faith in social engineering as such (Wickman, 1980). The report is a densely reasoned economic essay which put the notion of "planning" for social welfare firmly on the political map through emphasising the interactive complexity of social and economic explanations and the close integration of social, economic, legal, political and economic variables in reform preparation (SOU: 1934: 1). In a depressed economy, he argued, there was a need for increased public spending to stimulate recovery and "loan" funding for useful public works to improve employment, public assets and consumption, but also public morale. His main argument was that in view of the nature of business cycles, the "classical" strategy of maintaining a balanced budget, could be seen as a short term response, but was, in the long term, a counterproductive strategy. Instead the state budget should be balanced across cycles in order to offset deficits in lean years with surpluses gained in fat ones. An occasional unbalanced budget could be conceptually converted from being an unintended outcome of national economic policy, to an intended and managed occasional effect. Myrdal here clearly implies that with better information and planning we could correct what he saw as governmental "laid back" stupidities into *foreseeable*, rather than unintended, consequences. Such an approach would require not only longer term planning but also an ongoing empirical assessment of alternative "risk scenarios" for possible effects in a range of over lapping other domains such as international trade or monetary policy, the choice of which, given their hypothetical nature, was ultimately, he argued, a matter for political "judgement". What is seen as unintended, or unavoidable, is, given lack of information, only

unforeseeable. In this he spelled out the fluidity of intentions over time, and the difficulty of capturing them, but also the significance of better knowledge and information. He argued the importance of conceptual clarification of the "fluid" differences between various kinds of financial and social institutional regulations: hard core legal ones such as property and contract law, and more easily malleable ones such as rent and housing regulations. Altering the former, well entrenched in democratic law and culture, might appear desirable in the abstract, but not when assessed from the point of view of social stability and democratic traditions. Where to draw the line between them with respect to regulatory reform depends, he argued, on changing public attitudes and the relative power of interest groups and could not be understood as only a theoretical concern. What are required by the dynamic nature of the system requires, he argued, is a planned coordination of expectations and anticipations in the market as well as evidence of actual behaviour in the market and its real consequences.

Given that Myrdal saw the serious scientific issues in doing multidisciplinary analysis across different social domains as a problem with an enormous scope, he is left apologising for the "schematic" nature of his own analysis. In the an absence of data and information on core variables relevant to the making of predictions, such as the effects of income tax changes on consumption and investment, and much existing data too approximate and contested, there could be little basis for long term planning solidity. The solution to this uncertainty is seen by Myrdal to lie in what he saw as the sociologically well known, but seldom applied, methodological precept of a "comparative norm", i.e. a study of the difference between outcomes under one regulation, and outcomes under another. In practical politics, however, a comparative norm usually becomes defined as what would happen if a particular action is not taken. But then other actions, not part of the original model, are brought into play, thus invalidating the original comparison. For example, in assessing a particular kind of existing taxation it needs to be clear whether the effect is understood in comparison with what would happen if it was abolished concurrent with public expenditure reduction, or if such expenditure was maintained, and other methods of finance were employed such as state "loans". Because of such difficulties

in evaluating possible outcomes of alternative actions, he argued, his "presentation will be in some ways more abstract and difficult to grasp, than would be the case in a less conscientious report of the relevant economic presuppositions"(SOU: 1934: 1: 9). With some irony Myrdal here warns his readers that if they want him to be scientifically serious about social and economic planning, they should not expect an easy read or quick answers. On reception of this report however the Finance Minister was impressed about its quality and persuasiveness, and Myrdal's central position within the continued reform preparatory process was assured.[7]

Housing as a Model of Holistic Welfare Planning

One of Myrdal's next tasks was allocated to him by the Social Affairs Minister, who was convinced of the importance of the housing sector in an economic crisis policy. The arguments surrounding the development of a national housing policy, with the government both financially supporting and closer regulating the housing sector, became somewhat of a test case for what Holmwood refers to as the welfare state as guarantor of "complex freedom", in contrast to that of either a passive observer of market liberalism, or an authoritarian system of bureaucratic control (Holmwood, 2000). An investigative report was commissioned, with Myrdal as chair, on the need for the development of improved housing statistics (SOU 1933: 14). Prompted by increasingly desperate housing conditions during the depression and a previous lack of political will to act, improved housing was likely to gain voter support. It was also for Myrdal a topic which interestingly spanned overlapping economic and social domains, rich in variables and consequences to be managed. Existing studies of the rental market were not enough to give an indication of the latent need for new dwellings. This housing commission also had a multidisciplinary membership: an architect, a medical expert in bacteriology and hygiene, and a member of the social board which at the time dealt with issues of social aid and housing inspections. Like his previous assignment, this became a strategic "middle of the road" beginning in its rejection of state ownership of housing projects in favour of state investment and regulation, and an again apologetically "schematic" plan for the use of housing as an element in social engineering

across different social domains. It also had a multi-purpose: to promote knowledge about a public policy issue; to facilitate policy consensus on a politically contentious topic; and to act as a springboard for the development of improved planning research tools.

The city of Gothenburg, a major port known for its poor housing, was chosen as the location for a sample survey of housing needs and to provide some basic statistics regarding the relationship between overcrowding and income designed to inform both statistical improvements and housing reform. The report argued that if only relationships within the building sector could be regulated in a "sensible" manner, and planned with respect to its all pervasive consequences, the sector could contribute to reduction in unemployment and a boost in productivity. The commission's main dilemma is made clear from the start: there is an actual crisis situation with an urgent need of immediate public solutions, but there is also the problem generated by the need for long term planning requiring the development of better information, both local and national. A second dilemma is that "needs" must be determined from a social political perspective since credit support, both for individuals and for the building sector as a whole, leads to benefits for some, and potential losses for others through an artificial interference in the market, unless those in need are seen to be so poor as to be entirely outside the market. Therefore, if rational planning is to be possible, better housing statistics are necessary to establish the size and locations of the group with high needs but low purchasing power. When it comes to latent demand from groups with purchasing power, who could improve their housing condition, but choose to spend their income of other things, other measures may be necessary to force them to spend more on housing, such as improved information about the danger for children of overcrowding, or a strengthening of rules about housing standards. The "Leitmotif" which was to guide the Social Democratic reform process as a whole - an appeal to shared moral concerns for children, their health and development - here underlies the attempt to create consensus, but also concerns about fairness of outcomes. With individuals not interested in, or knowledgeable about, markets, and markets not considering the needs of individuals, more reliable information was needed to educate each side of the policy equation.

The report specified a long line of concepts in need of clarification such as "overcrowding", types of "latent need", minimum "housing standards" and different kinds of subsidy. Such measures, however, were not enough when it came to evaluating potential effects spilling over into other domains, particularly the institution of the family, previously, as the report notes, ignored out of traditional patriarchal concerns for the privacy of family life. It is repeatedly noted that government statistics have taken attention away from actual living conditions, to "surface" market phenomena such as the price of rents or the number of empty properties. What is needed is statistics on overcrowding in relation to family composition, standard of housing, family type, rent amount, social class, family income. Statistics should also be more inclusive of regional differences, using representative samples to avoid high costs, and the report argues for the need for a housing census to correspond with national population censuses.[8] But, projective statistics alone does not solve all planning problems. There are also issues about the nature of means employed. One noted potential unintended, but foreseeable, consequence of strengthened regulation is that of poor implementation of standards allowing for intentional malpractice or corruption in accounting or building practices. If the actual impact of reforms is not carefully and continuously measured, the extent of such corrupt practices cannot be established. In his later economic writings on welfare and development, both national and international, Myrdal regularly, to the chagrin of many politicians, returned to the issue of organisational and bureaucratic corruption, intended or unintended, in the implementation of what looked on paper like sensible rules (Myrdal, 1944; Myrdal, 1957; Myrdal, 1968; Jackson, 1990).

The longer the time span projected in a plan, the greater the number of unanticipated consequences and effects. The problem may seem insurmountable, but not impossible, argued Myrdal, as such programmes should not be rigid, but responsive to changes in external conditions, or to changing social and political perceptions of these conditions, all of which could be ascertained and studied with more effective social science methods less value loaded in favour of market interests. Long term programmes are needed to place a construction programme in its bigger context to ensure that short term perspectives

do not create hindrances for future developments. In the end Myrdal leaves us with a long list of questions for debate: How can one deal with the contradictory tendencies between demands for smaller or larger living units, all contingent on long term changes in cohort size, but also on externally induced changes in living standards, or changing behaviour with respect the production and rearing of children? The many questions raised in this report were left open for public discussion, and discussed they became in a variety of public and popular forums to which both the Myrdals and their architect friends eagerly contributed with their own particular conception of the welfare needs of the "modern" family in a "modern" productive nation (Olsson, 1990). But the complexities of social scientific "holistic" welfare planning had only just begun.

The Intended and Unintended Consequences of Birth Control

Two years after Gunnar Myrdal's *Spectrum* articles were published, the Myrdals' joint book, *Kris i Befolkningsfragan* ("Population Crisis") with its detailed proposals for state intervention in social politics reached the public (Myrdal and Myrdal, 1934). It was written, as Jackson describes it, with a rousing blend of "morality and dire prophecy" about the need for preventive, in their word "prophylactic", policy reform to stem a progressive population decline (Jackson, 1990: 76). Invest in families now, they argued, and the nation will guard itself against much greater possible costs in the future. Traditionally fertility debate had been dominated by conservative groups emphasising a continued ban on contraceptives, sex education and abortion in favour of a moral mainly religious approach to abstention and procreation. The Myrdals' joint book was the outcome of their efforts as young radicals to offer a multidisciplinary empirical synthesis of what they saw as wrong in traditional Swedish society, using fertility decline as the lynch pin around which to rally interest.[9] The book is rich in economic and sociological data and analysis, both statistical and more qualitative ethnographic descriptions of especially urban poverty and its consequences for families and children. Why would someone want to risk the cost of having children in such circumstances, they asked? But the sum total of seemingly rational individual decisions carried the unintended, or rather perhaps unforeseeable, consequences of population decline, a major long term risk to national

economic productivity and the financial and emotional support of individuals in families. Once foreseen, such an undesirable consequence could be turned around in that couples able to properly *choose* fertility in better social and economic circumstances might in fact be more likely to maintain population levels than those who could not. The book immediately became a bestseller, not least because of its open discussion of sex in and out of marriage and the actual spread of birth control as a factor in population decline, despite laws circumscribing its use, and the widespread use of illegal abortion. Such trends, they argued, could only be counteracted with greater social and economic security for women, most of whom wanted both work and children, and for children themselves as the future nation and its labour force. It is a nationalistic document in that it does not see immigration as a solution. It also strongly sociological, or in the term often used by Myrdal, socially environmental, and rejected firmly, using scientific genetic and well as social science evidence to firmly reject class and ethnic distinctions as a guide to whose children might be worthy of investment educationally and otherwise. Unlike the race biological arguments propagated by national socialist sympathisers, all children should be welcomed and all mothers supported. As Holmwood has argued, this book remains, alongside the work of Marshall and Polanyi, one of the most thoroughly argued statements in defence of state intervention in economic, social and especially family welfare (Holmwood, 2000).[10] Its publication broadened the parameters of what was seen as possible, but also necessary in preparing the country for a socially fairer and economically more productive national.

The intense public debate that followed in the press, parliament and study groups up and down the country, led to an agreement in 1935 between all major parties to establish a new commission on population issues. This offered Myrdal a fresh opportunity to exercise his social science expertise in the context of policy reform and raise the more practical methodological questions about what kinds of policies were likely to have the most beneficial effect for both individuals, seeking control over their own sexual relations and procreation, and the nation in need for a well cared for better educated and informed population (Lyon, 2001b). Myrdal was appointed secretary to the commission as a

whole and spent the next three years intensively working and writing primarily on the quality of its social science evidence and analysis. The nine formal members of the commission included members of parliament, varieties of medical experts, but also a professional gardener and a worker representative, as well as additionally, secretaries, assistants and co-opted experts. The team also included Alva Myrdal who became responsible for a report on married women and employment. When the final report of the commission was published in 1938, it had produced 20 sub-reports, 15 based on major social investigations (Wisselgren, 2006; SOU: 1938: 57).

The Population Commission was given a broad remit with respect to considerations about population quantity and quality and the impact of the overall situation of nativity decline leaving the long term support of the older generation in question. The sub-report of particular interest here, and the only one to be discussed, is the one devoted to fertility in relation to sexual behaviour, contraception and sex education. It became the most read of the overall commission's reports and sold 20,000 copies. Myrdal was the chair of this subcommittee and its brief included a critical report on the use of social science information and analysis in the issues involved. The population debate had proved a morally highly contentious one, and Myrdal took personal responsibility for writing on the methodological issues involved (SOU: 1936: 59). At his own insistence this was included in the final overall conclusion of the commission's work as a kind of public guarantee of research quality standards (Wisselgren, 2008).

Myrdal here emphasised the close and interactive relationship between various social domains: the economy, the family, inequalities in the education system, health and the labour market, and the short term and long term consequences across domains of ignoring recent fertility trends. He notes the absence of a factual discussion in a public debate characterised by lack of honesty in the use of evidence, especially as regards the use of contraceptives, and heavy moralising. Moralistic and medical debates, rather than informed arguments about the actual role of contraceptives within marriage and family formation have dominated debates. Society is, he argued, both too uninformed and misinformed to be able to make sensible reforms regarding fertility, especially

as information about advances in the technology of birth control could no longer be undone. With improved education and family supportive reforms, working together in mutual dependence and interaction, a more positive attitude to families and family formation would emerge.

In his methodological appendix, which offers a further example of the complexities in a social science based approach to policy reform. Myrdal starts by noting the lack of conceptual clarity in the discussion of population issues even in supposedly scientific presentations, especially when a fractional problem, such as prostitution or venereal disease, is discussed without its logical inclusion in what he saw as the larger question. In debates, he argued, special care has to be taken to clarify where ethical rather than factual judgements enter into debates and problem formulations. As population policy concerns the future, practical solutions to population issues must build on prognostic predictions about future trends, many of which are already given, and their causes and effects. But this raises a dilemma: as such predictions can never be perfect one must by necessity either operate with less than well grounded judgements, thereby easily falling into unstated ethical positions, or avoid prognosis altogether, thereby completely avoiding attempts at reaching grounded judgements. In this process of clarifying boundaries of uncertainty, the most important distinction to be drawn, he argued, is between *theoretical* and *practical* problem formulations, between the description and explanation of social reality in terms of cause and effect, and the valuations that determine the choice between possible alternative actions in response to a theoretically described and causally explained reality. This distinction, he argued, is in practice difficult to maintain. Both sets of problem formulations have a logical connection, and within both there is a *qualitative* and *quantitative* side. The first concerns changing physical, psychological and moral population characteristics. The second concerns its size, changes and trends with respect to variables such as migration, age, marriage, and birth rates. There is a complex relationship between these two sets of variables in that the quantitative aspects of population size and structure are ultimately determined by the qualitative ones. But the reverse is also true as quantity indirectly has an impact on population quality, a conclusion drawn from his earlier work on housing standards and its consequences

for children, families and more broadly education and the work force in general. In this interactive dynamic it is difficult to disentangle what is cause and effect making attempts at prognosis highly scientifically speculative.

In terms of the applied practice of drawing up policy programmes, qualitative and quantitative aspects also run together. If factual analyses of the present are combined with different value priorities, including those relating to other potentially affected domains such as the labour market, or child health, a great variety of possible practical conclusions emerge. Analysis of past causes and effects can be undertaken, using as firm an empirical base as possible. But a causal analysis of variable relations in the past, however well done, can only tentatively be stretched into the future as prognosis due to limited information about so far unknown extraneous future events, or changes in social values. The longer into the future a prognosis aims to project, the more uncertain it becomes. When it comes to indirect pro-creational motivational causes, which are likely to affect the psychology of choice, these can only be projected through a total social institutional and structural analysis. In such an analysis any thought of finding one overriding cause of nativity change disappears, 'unless one strives to find one primary cause for the totality of social development and then one is in the domain of social metaphysics' (SOU: 1936: 59: 152). 'Social metaphysics' was a term Myrdal tended to use when referring to religious and other strongly ideological positions of unwarranted predictive certainty. Between an individual's understanding and behaviour there is a host of social institutions, laws, conventions, group habits, action patterns and social norms that both determine and are determined by the technical and material developments and by the reactions of other individuals. Myrdal concludes that a causal discussion of the nativity problem becomes useless if it is not seen as part of a total sociological perspective and empirically treated within the frame of particularly family sociology.

With the problem of causality altering relative to a particular time, the causal analysis "of the moment" usually looks only at individual *motives* and intentions vis a vis birth control and sexuality more abstractly. Such interpretations assume the actions are "conscious and intentional", which, he argues, evidence shows to be both factually false,

as individuals hold incorrect assumptions about birth control, and value oriented in that individuals relate to other values than those pertaining to the use of birth control alone. Judgement of birth control issues thus becomes, he argued, one of expressing views of what "hypothetically" individuals should do if they knew properly, or if they adopted the "right" norms, a kind of constructed "standard motivation" not very relevant to how and what individuals think as real individuals. With the contexts of choice and motivation changing over time, decisions are thus not entirely determined by reason, but by time and location. Many "motives" offered as answers to questions in various forms of social research are thus "post rationalisations" and needs to be treated with a great deal of scientific caution.[11] Besides, different kinds of motives are woven into each other with little distinction made between factual, moral and other more practical motives. If one could scientifically disentangle this, the problem would be easier to explain. Myrdal's own enlightenment faith here comes to the fore when he argues that many differences of opinion would dissolve, if such confusions could be more apparent and clearly stated.

The magnitude of difficulties in trying to establish trends and causalities in population change are further increased when problem formulations are broadened into practical and socially goal oriented ones. A policy conclusion becomes hypothetical in a different way than that of a sociological prognosis. There are political elements in the norms postulated when individual motives are judged, each framed by different interests, and when means and goals are at stake. Though policy goals cannot sensibly be set other than on the basis of as real as possible knowledge of different possible population projections, in the final instance the desirability of certain goals and the means to achieve them are subjectively differentiated relative to political interests. Such valuations ought to be made conscious and brought in as explicit premises for conclusions drawn. In a policy programme value based interventions become new causes to be investigated with respect to their possible effects.

The final dilemma of complexity is both a technical and a moral one: if the investigation should extend into all possible alternative future actions and their possible consequences it would be boundless and incalculable. Looking at factual probabilities based on past developments several possibilities can however be excluded. From a theoretical viewpoint,

a multitude of valuations are also hypothetically possible. But when it comes to values, however, only those means and goals ought to be included which have enough of a basis in people's value consciousness to be of political significance. Only certain value positions are socio-politically real enough as goal orientations amongst citizens and groups of citizens to the extent that it can be motivated to use them as foundation for a practical analysis. For Myrdal this proved a value basis on which to support the decision of the Social Democratic Party to rule out any attempts at speeding up a process of nativity growth through for example 'paying for babies' and keeping women in the home, policies introduced in Germany at the time, but seen as undemocratic by the final report of the commission. Instead measures such as improved housing, support for child care, support for children born out of wedlock, protection of married women's employment, 'marriage loans', improved health care for children, including support for the transport of fresh fruit and vegetables to the Northern districts, were instigated. Sweden's reputation as the very model of a welfare state was born.[12]

CONCLUSION

Two decades later, in a paper to the inaugural conference of the British Sociological Association, Myrdal tried to sum up his Swedish and international experiences as a 'working' social scientist and make some predictions about the future for social scientists in policy making. By then he had managed and written the final report for one of the largest and most definitive social science projects ever undertaken: on the topic of racism in the US and its many "odious" intended and unintended, consequences (Myrdal, 1944; Lyon, 2004). He had also for a period shortly after the war been director of the Economic Commission for Europe, the first European regional organisation created by the UN to initiate measures for economic reconstruction and integration (Barber, 2008). Both were heavily value and conflict laden tasks for which he got little thanks and saw few positive outcomes in a world dominated by the cold war. Welfare state bureaucracies were rapidly expanding. His methodological work had given birth to the dynamic complexity-catching concepts for which he is best known: 'cumulative causation' and 'vicious

and virtuous circles' of social change (Barber, 2008). These were, in his terms, 'do something' rather than 'do nothing' concepts in causal thinking (Myrdal, 1970). His optimism about the 'romance of the engineer' and the effectiveness of rational debate in politics had abated, but the problem of how to deal with contradictory value premises in human personal and political actions remained, a problem he continued to believe should not be hidden away behind complex methodological technologies, thereby deceiving those whose lives are affected by them. In reality, he argued, there is no certainty in social prediction and 'a desired end, if reached, is never attained in purity, but the dynamic process initiated by the means results in many other changes besides the positive achievement of the end' (Myrdal, 1953: 239).

Science grows by enlightened criticism and social science imply criticism of society, its values and the immediate and long term consequences of social and political actions. But state institutions, he argued, are not likely to be supportive of value critical social science. Politicians' only interests are in the present and they must compromise on truth to remain in power. Civil servants are essentially servants of power and must care about detail and routine rather than long term trends. Journalists cater for the market interest in day to day affairs. We can perhaps, in the light of the arguments presented above, rephrase these questions as: who will be there to hypothesise about consequences, turn unintended consequences into foreseen and manageable ones, variable complexity into planned outcomes for society as a whole, value premises into democratically supported strategic action? Will increased policy importance impair academic freedom also within universities which has traditionally sheltered the growth of social science? The dangers outlined relate both to personal and institutional morality with temptations into corruption created by growing pecuniary awards and status advancement, and by demands for loyalty to institutions, including governments, rather than truth and critical public scrutiny of evidence collected and value interests pursued. Half a century later and in the midst of yet another socially destructive economic crisis with disastrous welfare consequences that should, with better social science understanding, have been at least foreseen if not prevented, it is hard not to share his pessimism. The fact that sociology today remains a discipline

that regards itself as somehow in the service of humanity, yet does not see itself as entitled to take a stand in support of sections of humanity for fear of being seen as "biased" or "unscientific", remains a troubling matter (Flyvbjerg. 2001; Sayer, 2011). We would do well to share, and propagate, Myrdal's faith in the necessity of social science as a tool in educating both politicians and the public about actual states of affairs and in clarifying the complex relationship between values we pursue and their outcomes lest we abandon the very persons about whom we invariably continue theoretically to generalise.

NOTES

1. The work here discussed was aimed at social policy reform. To avoid confusion with the academic discipline of social policy, which shares many theoretical and methodological perspectives with sociology but tend to focus more on policy administrative processes, general discussion about social policy will be referred to as "social politics", a term commonly used in Swedish.
2. The close personal, intellectual and working relationship between the Myrdals have been much scrutinised (see for example Jackson, 1990; Hirdman, 2008; Ezemuller, 2010). For present purposes, discussions will be attributed according to formal authorship only, without any attempt at holding either responsible for particular views.
3. Note on Ogden.
4. Myrdal did not here emphasise the meaning of "Enlightenment", a concept to which he often proclaimed his allegiance. Like other enlightenment thinkers its three major tenets of individual autonomy, the universality of humanity itself as the end of purposive action, and the search for knowledge, freely reasoned and sought, to support this, underlay much of his thought on science as well as politics (Appelquist *et al*, (eds), 1998; Todorov, 2010). On his insistence it was the last word on which his major work on racism in the US ended (Myrdal, 1944; Barber, 2008: 76).
5. Here Myrdal outlines a theoretical, political and methodological "middle way" forward. After 1945, when the Swedish welfare system became generally known as a "third way" model, this came to refer more to a system of economic bargaining between the state, capital and trade unions. But the source of the model, as Myrdal's theoretical synthesis, similarly had some of its origins in pre-National Socialist German Social Democratic Party history (Ekdahl, 2001).
6. This is the topic for which Myrdal received the Nobel Prize in Economics in 1974. He shared this with von Hayek, who started his economics career in the same field. They were to follow very different and opposite intellectual paths later in life with Hayek regarding government intervention as the source of most social and economic problems.

7. There is ongoing debate in economics about the relationship between Keynesian economics and that of Myrdal and the Stockholm school, partly caused by the fact that the works of the latter did not reach, nor were they intended to reach, an international audience at the time (Barber, 2008). The idea of state planning and public investment to boost economies was, like social engineering, an idea "in the air" at the time.
8. Such statistics and recording systems are today part and parcel of everyday administrative practices of the civil service and local authorities. They are still, however, subject to regular political debate and even political interference to boost particular policy plans.
9. Lest it be argued that the adoption of fertility issues was a cynical ploy to gain support for their views, Gunner Myrdal's more explicitly political writings in party pamphlets indicate his passionate personal belief in the right of individual men and women to control over their own sexual and procreative lives (Myrdal, 1936). Alva Myrdal, who was pregnant at the time of their joint writing on the fertility question, saw the humour in the situation when she wrote that her psychiatrist friend had laughed at them "for being so devilishly intellectualized that we have to write three hundred pages in order to clear out our motives behind the decision to multiply" (Andersson, 1999).
10. The role of the Myrdals and their belief in state intervention has been much discussed in the context of a debate about sterilisation policy in Sweden after the war. As has now been clarified, they did not in their writing advocate involuntary sterilisation except in very exceptional circumstances. The high number of "voluntary" sterilisations that took place after the war is in the light of later evidence to be attributed either to "tired" older mothers not wanting more children, or to the activities of moralistically overzealous medical practitioners. This debate is in itself an interesting study in the power of unintended consequences and the need for real social evidence in the attribution of blame for social ills (Broberg and Roll-Hansen (eds) 1996; Ekerwald, 2001).
11. This is a theme Myrdal later developed in his theory of the structure of racist attitude and beliefs and explains his critical view of attitude surveys (Myrdal, 1944). It bears some similarities with C. Wright Mill's conception of "socially situated motives" (Mills, 1940).
12. The speed with which the early Swedish welfare reforms with their sexually liberal assumptions and their family and women friendly focus reached an international audience can be attributed to the publication in the US in 1941 of Alva Myrdal's still much read book on the topic (Myrdal, A., 1941).

REFERENCES

Appelqvist, O. and Andersson, S. (eds) (1998) *Vagvisare: Texter av Gunnar Myrdal.* Stockholm: Nordstedts Forlag.

Andersson, S. (1999) "On the value of personal archives", *Nordeuropa Forum,* 1: 15–32.

Barber, W. J. (2008) *Gunnar Myrdal*. Basingstoke, Hampshire: Palgrave Macmillan (Great Thinkers in Economics Series).

Berggren, H. and Tradgardh, L. (2009) *Ar svensken manniska? Gemenskap och oberoende i det moderna Sverige*. Stockholm: Nordstedts Forlag.

Broberg, G. and Roll-Hansen, R. (eds) (1996) *Eugenics and the Welfare State: Sterilization Policy in Norway, Sweden, Denmark and Finland*. East Lancing: Michigan State University Press.

Burawoy, M. (2005) "2004 American Sociological Association Presidential Address: 'For Public Sociology'", *The British Journal of Sociology*, 56(2): 259–94.

Carlson, A. (2005) *Fractured Generation*. New Brunswick, New Jersey: Transaction Publishers.

Cherkaoui, M. (2007) *Good Intentions: Max Weber and the Paradox of Unintended Consequences*. Oxford: The Bardwell Press.

——— (2008) *Durkheim and the Puzzle of Social Complexity*. Oxford: Bardwell Press.

Ekdahl, L. (2001) *Mot en tredje vag: En biografi over Rudolf Meidner—Del 1* (Towards a Third Way: A Biography of Rudolf Meidner - Part 1). Lund: Arkiv Forlag.

Ekerwald, H. (2001) "The modernist manifesto of Alva and Gunnar Myrdal: Modernisation of Sweden in the thirties and the question of sterilisation", *International Journal of Politics, Culture and Society,* 14(3): 539–561.

Eliason, S. (2000) "Gunnar Myrdal: A Theorist of Modernity". *Acta Sociologica*, 43(4): 331–341.

Eyerman, R. (1985) "Rationalising intellectuals: Sweden in the 1930s and 1940s", *Theory and Society,* 14(6): 777–807.

Ezemuller, T. (2010) *Die Romantik der Rationalitat: Alva and Gunnar Myrdal– Social Engineering in Schweden*. Bielefeld: transcript Verlag.

Flyvbjerg, B. (2001) *Making Social Science Matter: Why Social Science Fails and How it Can Succeed Again*. Cambridge: Cambridge University Press

Giddens, A. (1998) *The Third Way: The Renewal of Social Democracy*. Cambridge: Polity Press.

Hirdman, Y. (2008) *Alva Myrdal: The Passionate Mind*. Bloomington: Indiana University Press

Holmwood, J. (2000) "Three pillars of welfare state theory: T. H. Marshall, Karl Polanyi and Alva Myrdal in defence of the national welfare state", *European Journal of Social Theory*, 3(1): 23–50.

Honegger, C. and Wobbe, T. (1998) "Einleitung: Frauen in der kognitiven und institutionellen Tradition der Soziologi" (Introduction: Women in the Cognitive and Institutional Tradition of Sociology), in C. Honegger and Wobbe, T. (eds.) *Frauen in der Soziologie: Neun Portrats* (Women in Sociology: Nine Portraits). Munich: Verlag C. H. Beck, pp. 7–27.

Jackson, W. A. (1990) *Gunnar Myrdal and America's Conscience: Social Engineering and Radical Liberalism 1938–1987.* Chapel Hill: University of North Carolina Press.

Kaplan, A. (1964) *The Conduct of Inquiry: Methodology of the Behavioural Sciences*. San Fransisco: Chandler Pub. Co.

Lepenies, W. (1988) *Between Literature and Science: The Rise of Sociology*. Cambridge: Cambridge University Press.

Lyon, E. S. (2001a) "The Myrdals and the Thomases: the trials and tribulations of a cross-Atlantic research collaboration", in Mucha, J., Kaesler D. and Winclawski, W. (eds), *Mirrors and Windows: Essays in the History of Sociology*. Torun: Nicholas Copernicus University Press/International sociological Association.

——— (2001b) "Education for modernity: the impact of American social science on Alva and Gunnar Myrdal and the 'Swedish Model' of school reform", *International Journal of Politics, Culture and Society,* 14(3): 513–537.

——— (2004) "Researching race relations: Myrdal's American dilemma from a methodological perspective", *Acta Sociologica,* 47(3): 203–217.

Mannheim, K. (1972) *Ideology and Utopia: An Introduction to the Sociology of Knowledge*. London: Routledge and Kegan Paul Ltd. (revised version of *Ideologie und Utopie* first published in Germany in 1929).

McLelland, G. (2004) "Travelling with vehicular ideas: the case of the Third Way", *Economy and Society,* 33(4): 484–499.

Merton, R. K. (1936) "The unanticipated consequences of purposive social action", *American Sociological Review,* 1(6): 894–904.

——— (1957) *Social Theory and Social Structure*. New York: The Free Press.

Mills, C. W. (1940) "Situated actions and vocabularies of motive", *American Sociological Review,* 5: 439–452.

Myrdal, A. (1941) *Nation and Family: the Swedish Experiment in Democratic Family and Population Policy*. New York: Harper and Brothers.

Myrdal, A. and Myrdal, G. (1934) *Kris is Befolkningsfragan*. Stockholm: Tiden.

Myrdal, G. (1932) "Social politikens dilemma" (The dilemma of social politics), *Spektrum* 2(3): 1–13 and (4): 13–31.

——— (1936) *Vad galler striden i befolkningsfragan?* Stockholm: Frihets Forlag.

——— (1944) *An American Dilemma: The Negro Problem and American Democracy*. New York: Harper and Brothers (3rd edition 1996. New Brunswick, NJ: Transaction).

——— (1953a) *The Political Element in the Development of Economic Theory*. London: Routledge and Kegan Paul (Published in Swedish in 1930 and in German in 1932).

——— (1953b) "The relation between social theory and social policy", *British Journal of Sociology,* 4(3): 210–242.

——— (1957) *Beyond the Welfare State: Economic Planning in the Welfare States and its International Implications*. London: University Paperbacks.

——— (1968) *Asian Drama: An Inquiry into the Poverty of Nations*. London: Allen Lane.

——— (1970) *Objectivity in Social Research*. London: Gerald Duckworth and Co. Ltd.

Olsson, S. E. (1990) *Social Policy and Welfare State in Swede*. Lund, Sweden: Arkiv Forlag.

Oppenheimer, F. (1962) *Weder Kapitalismus Noch Kommunismus* (Neither Capitalism nor Communism). Stuttgart: Gustav Fischer Verlag (first published in 1919).

Podgorecki, A., J. Alexander and R. Shields, eds. (1996) *Social Engineering*. Carleton: Carleton University Press.

Sayer, A. (2011) *Why Things Matter to People: Social Science, Values and Ethical Life*. Cambridge: Cambridge University Press.

Smelser, N. J. (1997) *Problematics of Sociology: The George Simmel Lectures, 1995*. London: University of California Press.

SOU: 1933: 14 *Bostadsfragan som social planlaggningsproblem: Under kris och pa langre sikt. En undersokning rorande behovet av en utvidgning av bostadsstatestiken jamte cvissa darmed forbundna bostadspolitiska fragor.*

——— 1934: 1 *Finanspolitikens ekonomiska verkningar.*

——— 1936: 59 *Betankande i sexual fragan.*

——— 1938: 57 *Slutbetankande angivet av Befolkningskommissionen.*

Todorov, T. (2010) *In Defence of the Enlightenment*. London: Atlantic Books.

Turner, S. P. (2007) "Public sociology and democratic theory", *Sociology,* 41(3): 785–797.

Wagner, P. (2001) *A History and Theory of the Social Sciences: Not All that is Solid Melts into Gold*. London: Sage.

Wickman, K. (1980) *Macro-ekonomisk planering—orsaker och utveckling* (Macro-economic Planning: Origins and Development). Uppsala, Sweden: Almquist and Wiksell International.

Wisselgren, P. (2006) *Regulating the Science Policy Boundary: The Myrdals and the Swedish Tradition of Governmental Commissions*. Uppsala University: Stella Arbetsrapporter, Nr. 31.

4

"ORIENTATION TO OTHERS"

A Central but Forgotten Concept in Max Weber's Sociology

Richard Swedberg
Cornell University

One of my favorite writings of Mohamed Cherkaoui is his book *Good Intentions: Max Weber and the Paradox of Unintended Consequences* (2007).[1] The main reason is that this study mixes in a truly ingenious way an attempt to use the ideas of Max Weber with an interest in social mechanisms. Following Mohamed Cherkaoui, and in homage to his work, I have tried to follow this strategy in this brief article.

In the first sentence of §1 of Part I in *Economy and Society* Weber presents his well-known definition of sociology. Sociology, the reader is told, is a science that focuses on "the interpretive understanding of social action". In the next two sentences he explains that what he means by "action" and "social". Action is defined as behavior to which the actor attributes "subjective meaning". What makes action "social" is that "its subjective meaning takes account of the behavior of others and is thereby oriented in its course" (Weber, [1922] 1978: 4).

In this paper I will argue that one part of this definition has largely passed unnoticed, but is nonetheless central to Weber's view of sociology. It is the definition of what is "social" and it can be found in the

last sentence of §1 that deals with action being "*orientated in its course* [*to others*]".

The last sentence of §1 of Part I *in Economy and Society* reads as follows in English: "Action is 'social' insofar as its subjective meaning takes account of the behavior of others and is thereby oriented in its course". This is a translation of the following sentence: "'Soziales' Handeln aber soll ein solches Handeln heissen, welches seinem von dem oder den handelnden gemeinten Sinn nach auf das Verhalten *anderer* bezogen wird und daran in seinem Ablauf orientiert ist" (Weber ([1922] 1972: 1).[2]

The secondary literature devoted to Weber's famous definition of sociology in §1 of *Economy and Society* is large and has especially tried to establish what Weber means with *Verstehen* and meaning. Practically no attention, however, has been devoted to the fact that Weber qualifies his statement of what constitutes the general subject of sociology (meaningful behavior) by saying that it also has to be social or *oriented to the behavior of others*.

This is all the more peculiar since it is precisely this element that separates out sociology from the other cultural sciences and provides it with its very own subject matter. While many of the cultural sciences deal with meaningful behavior (such as history, political science, economics and so on), what defines sociology as a separate science is precisely that it studies meaningful behavior ("action") that I oriented to others ("social").

In this paper I will try to rectify this omission by focusing squarely on the concept of orientation to others in Weber's sociology (Part I). I will also introduce a few examples from works in social science in which a very similar concept has been used, in order to illustrate the fruitfulness and flexibility of Weber's concept (Part II). I shall conclude by discussing which aspects of "orientation to others" still remain to be explored; how this concept differs from the concept of social interaction; and how it may be further improved (Part III).

I. WEBER'S CONCEPT ORIENTATION TO OTHERS

Given the fact that the concept of *orientation to others* is central to Weber's definition of sociology, a number of questions arise. How, for

example, does Weber define "orientation to others", and how does he use it in his sociology? Did Weber borrow this concept from somebody else's work or did he invent it himself? And, finally, when does the concept of orientation to others first appear in Weber's work and how does it evolve over time?

When confronted with questions of this type, one naturally turns to the enormous secondary literature on Weber's work which has accumulated over the years. In this particular case, however, very little is to be found.[3] In his famous *Structure of Social Action* (1937) Talcott Parsons, for example, cites Weber's definition of sociology in *Economy and Society* and also comments on the various elements of Weber's definition of sociology (Parsons, [1937] 1967: 641, 649).[4] He, however, has nothing to say about the element of "orientation to others".

Pretty much the same is true for other standard works on Weber's sociology by such well-known scholars as Raymond Aron ([1967] 1970), Julien Freund ([1966] 1972) and Alexander von Schelting (1934). The main thrust in these works is that the two key elements in Weber's definition of sociology are "meaning" ("*Verstehen*") and "action". That these two elements also are constitutive for other social sciences ("cultural sciences") than sociology, especially history and economics, is not commented upon.

Of the major interpreters of Weber, there is to my knowledge only one person who has paid attention to the fact that orientation to others represents a distinct concept in Weber's work and that it is constitutive for the way in which he conceptualizes the very nature of sociology. This is Alfred Schutz in *The Phenomenology of the Social World* (1932).

Schutz' interpretation, however, is somewhat limited for the reason that he basically uses Weber's concept of orientation to others as a prelude to his reworking of this concept, for theoretical purposes of his own. He is not, in other words, very interested in how Weber himself understood and used the concept of orientation to others. Nonetheless, Schutz does make some interesting comments on Weber's concept of orientation to others, and I shall return to these in a moment.

In the rest of the secondary literature on Weber one can find an occasional comment or two on the notion of orientation to others, but nothing that is either comprehensive or particularly penetrating.[5] No

author singles out orientation to others as a special concept in its own right, and it therefore remains at the level of what Robert K. Merton has called a *proto-concept*, that is, an early and underdeveloped version of what may later become a scientific concept.[6] As the reader soon shall see, I have however found Dirk Käsler's (1988) discussion quite helpful.

It is not known whether Weber himself invented the concept of orientation to others or borrowed it from somebody else's work, as he so often did. If he borrowed it from somebody else, a good candidate would be Kant. The German philosopher used the concept of orientation in many of his writings and especially in a well-known article called "What is Orientation in Thinking?" ([1786] 1991).

But Weber typically changed the concepts that he borrowed from the work of others in a fundamental manner, as exemplified by the notions of charisma and the ideal type. One should therefore not expect to find the key or "explanation" to what he meant by some term by looking at what may well have been its source of inspiration. This would also seem to be the case with "orientation", at least in so far as Kant is concerned.[7]

It is also possible today, through various electronic editions of Weber's work, to establish when he first used a term, how frequently he used it and the like. Since this paper, however, deals with Weber's use of his concept in his sociology and not in his pre-sociological works, I will leave this task to others.

What I will focus on in this article is primarily the role of "orientation to others" in the crucial *Economy and Society* and the writings that Weber wrote in preparation for this work. Of particular interest are especially the following two writings: the article that was published in *Logos* in 1913, and Part 1 of Volume 1 in *Economy and Society*. The reason why so much importance should be attached to the *Logos* article ("Some Categories of Interpretive Sociology") is that it represents Weber's first major attempt to establish a theoretical foundation for sociology as a special science in its own right.

Of particular interest is also Part 1 in Volume 1 of *Economy and Society*, since it represents the only section of the book that we today know as *Economy and Society* which Weber himself cleared for publication (e.g. Schluchter, 1998). It is also in Chapter 1 of Part 1 that we find

the final version of Weber's attempt to establish a theoretical foundation for sociology.

Weber's article in *Logos* is very complex and deserves a much fuller discussion than it can be accorded here. Nonetheless, it shall be pointed out that the concept of orientation to others has roughly been assigned the same place in the definition of sociology that can be found in the *Logos* article as in *Economy and Society*. That this is the case is clear from the following central quotation about the task of sociology:

> That kind of action that is specifically important for interpretive sociology is, in particular, behavior (1) that, in terms of the subjectively meant meaning of the acting person, is related to the *behavior of others*; (2) whose course is *partially determined* by this meaning-relatedness; and (3) that can therefore, on the basis of this (subjectively) meant meaning, be intelligibly *explained*. (Weber, [1913] 2012: 274; cf. Weber, [1913] 1981: 152)

As the reader may have noticed, Weber uses the term "related to", rather than "oriented to", in the definition just cited. This particular term, however, is used by Weber as synonymous with "oriented to", and it can, for example, also be found in the definition of sociology in Chapter 1 in Part 1 of *Economy and Society*.[8] It should be added that while Weber does not use the term "orientation to" in his definition, it is frequently used throughout the article.

It is, however, in Part 1 of *Economy and Society* that we find Weber's fullest and most official explication of the concept of orientation to others. The great part of what Weber has to say on orientation to others can be found in Chapter 1 (general sociology) and in Chapter 2 (economic sociology); in Chapter 3 (domination) and Chapter 4 (status groups and classes) he only makes an occasional use of this concept. Since this paper is focused on general sociology, I will primarily be drawing on Chapter 1, and especially on the section in this chapter which is exclusively devoted to the concept of orientation to others (Weber, [1922] 1978: 22–4).

In his chapter on general sociology in *Economy and Society* Weber starts out by carefully outlining what kind of actions do *not* fulfill the criteria of being oriented to others, and which consequently do not

qualify as "social action". There is, first of all, the demand that the action must be oriented in a meaningful way to what someone else does. A bicycle rider, for example, who collides with another bicycle rider, has not oriented his or her behavior to the other person; they just happen to bump into each other by accident.

It is important to point out at this stage of the discussion that just as Weber's interpretive sociology is centered around the idea of *meaning* (*Sinn*) or human behavior that is understandable, so orientation has an element of meaning to it. At one point in his explication in Paragraph 1 of what constitutes "social behavior" (which is the section where the concept of orientation is explicated in Paragraph 1), Weber describes social action as being "*meaningfully oriented*" (Weber, [1922] 1978: 23; emphasis added). An act, to give a concrete example, may have a different meaning if it is oriented towards someone who is, say, a Protestant than a Catholic.

The action, Weber continues in his explication of the sociological use of the concept of orientation, also has to directed at another *person.* Solitary prayer or several people who simultaneously open their umbrellas when it starts to rain, do not constitute (meaningful) actions that are directed at other people—and hence do not qualify as social actions. Actions which are exclusively oriented to values and facts constitute further examples of actions which are not to be characterized as social action; and so do actions oriented to inanimate objects.

But even if one keeps all of these qualifications in mind, the area which Weber assigns to "social action" is very large, and it includes a great number of different types of action. To give a proper account of how Weber views social action, one would have to go through Chapter 1 paragraph by paragraph and summarize what Weber says in each of these. Given the size and the aim of this paper, which is to draw attention to Weber's concept of orientation to others, this way of proceeding is not appropriate. From a more general perspective, however, one can say that Weber analyzes different types of social action according to *time,* to *the object* toward which they are oriented, and to *the way* they are oriented; and a few words will be said about each of these categories.

As to time, Weber notes that social actions can be oriented to the past, to the present, and to the future. Social actions can also be

oriented towards individuals, and towards what Weber calls "orders". Social actions, finally, can be oriented towards individuals and orders in four major ways: in an instrumentally rational manner, in a value-rational manner, in an affectual manner, and in a traditional manner.

Little need to be said about the time dimension of social action and that social action can be directed towards the past, the present and the future. As an example of an act which is oriented to the past, Weber mentions revenge. There also exists an enormous secondary literature on the four different ways in which social action can be oriented (instrumentally rational, value-rational, affectual and traditional action). The only thing that needs to be added here is that Weber states that these four ways of acting do *not* constitute social action per se; for this, they also have to be meaningfully oriented at the behavior of others. This last point is rarely, if ever, mentioned in the secondary literature. Still, it vitally affects our understanding of these types of action, especially value-rational action.[9]

The objects of social actions are not only individuals but also orders, as we learn as we continue to read *Economy and Society*. The three basic configurations that result from Weber's argument have been helpfully summarized by Dirk Käsler (1988: 157). There is first of all the single individual who may orient his or her actions to another individual or to many other individuals. You may, for example, orient yourself to another person by addressing him or her, by looking at him or her, and so on.

The second configuration is that of mutual orientation, between two or more individuals. Examples of mutual orientation include such well-known Weberian concepts as struggle, competition, power and domination (see Weber, [1922] 1978: 38–40, 53–4). What Weber calls associative and communal relationships fall in this category as well.

Actions can also—and this is the third configuration—be oriented towards so-called orders. An order (*Ordnung*), Weber explains, originates in a social relationship, and is characterized by the fact that its content can be summed up in a few maxims of behavior (Weber, [1922] 1978: 28, 31; cf. Swedberg, 2005). These maxims appear as exemplary or obligatory to the actors. Organizations, it should be noted, constitute a special kind of order (closed social relationships which are enforced by a staff).

It is clear that there exist combinations of these three categories. A person may, for example, orient his or her actions to another person and an order at the same time. Two orders can interact through the agency of two individuals, and so on.

Weber discusses some of these combinations in *Economy and Society*. He also notes that an act can be oriented to more than one order simultaneously, and gives as an example someone who decides to duel, despite the fact that dueling is forbidden. He further notes that being oriented to an order does not necessarily means that one accepts its maxims. A thief, for example, orients his or her actions towards the law, but also attempts to outwit its guardians.

It would also seem clear that the concept of orientation lends itself in some cases to a networks analysis. I will not pursue this issue here, however, but instead proceed to the next task, namely a critical discussion of the way that Weber constructs and makes use of the concept of orientation to others in his sociology.

First of all, it has to be admitted that there is something to Alfred Schutz' charge that Weber's concept is not very clear (Schutz, [1932] 1967: 17). For one thing, Weber confuses things by using the word "orientation" in several different meanings. One example of this is when Weber in his economic sociology introduces the term "economically oriented action", where the word orientation is used in a different way (Weber, [1922] 1978: 64–8).[10]

Weber also states very clearly that social action must be oriented to "the behavior of others" in his definition of sociology in Chapter 1 in *Economy and Society*—only to later say that it also can be directed at an order. It is finally also true that Weber nowhere in *Economy and Society* (or elsewhere) provides a definition of what he means by the expression "orientation to". How, for example, does this concept differ from the psychological concept of perception? And how is it related to symbols and the use of language?

Schutz is also very critical of Weber for not properly distinguishing between situations in which an actor affects someone else, and situations in which an actor is affected by someone else (Schutz, [1932] 1967: 144ff.). It is not clear to me what consequences this distinction has for Weber's concept of orientation to others, and this also goes for some related points that Schutz makes in *The Phenomenology of the Social World*.

Schutz does not spend much time on explaining these things, and the reason for this, it seems to me, is that he is primarily interested in criticizing Weber's concept of orientation to others for theoretical purposes of his own. What is beyond doubt, however, is that Schutz should be commended for having realized that orientation to others plays a crucial role in Weber's sociology. More work of similarly penetrating character is clearly much needed.

But even if there are some problems with Weber's concept of orientation to others, it is also important to realize how much Weber actually accomplishes with the help of this concept. First and foremost, it allows him to recast a good part of standard sociology and produce fresh and interesting definitions of concepts such as conflict, competition, power, and so on.

The idea of orientation to others also goes well with Weber's view of the role of contingency in social life. Actors often orient themselves to *expectations* about what other actors will do, based on their earlier experiences, Weber says, but these expectations may not coincide with how these other actors will later behave. There consequently exists a strong link between Weber's concept of orientation to others and his concept of probability (*Chance*).

The concept of orientation to others also fits very well with Weber's vision of society as an ongoing process. Society, in Weber's mind, does not consist of collective actors and solid, objective structures, but of individuals who orient their actions to what they think other actors will do.[11] There is also the fact that an additional amount of flexibility (as well as complexity) is introduced into the analysis through Weber's idea that an actor can simultaneously be oriented to several orders and/or actors. All in all, one can say that Weber's concept of orientation to others is very flexible and central to his general view of sociology.

II. LATER AND RELATED USES OF THE CONCEPT OF ORIENTATION TO OTHERS

But even if Weber makes good use of the notion of orientation to others in *Economy and Society*, is it also a concept that more generally can be of help in sociology? In trying to answer this question I shall refer

to some research by other sociologists and social scientists than Weber who on their own have made use of the idea of orientation to others without necessarily using this exact formulation. I shall first say something about the way that G. H. Mead made use of this idea, then about its role in reference group theory, and finally I will comment on the way that it has been used by economists, in particular by Thomas Schelling. In each of these examples I shall try to establish both the centrality of the idea of orientation to others and also how a particular study adds to Weber's original use.[12]

G. H. Mead and Weber conceived their most major ideas around the same time, but do not appear to have been aware of one another (e.g. Joas, 1985: 90–102). That the notion of orientation to others is central to Mead's work is immediately clear from such concepts as the social act and taking the role or attitude of the other. In taking the attitude of the other, for example, the actor needs to figure out what the other person thinks, and adjust his or her behavior to this. It should also be pointed out that this form of orientation and action can be carried out in a more or less deliberate manner. If Actor A wants to act in a rational way, he or she will try to figure out the possible strategies of Actor B, and then adjust his or her behavior to these. Game theory, as I shall return to, can be seen as an attempt to analyze precisely this type of rational behavior.

But it may also be the case that *all* forms of social action (in Weber's sense) hinges on the capacity of actors to see things from the perspective of other actors. In order to orient oneself to someone else's behavior in a meaningful way, both actors have so-to-speak to inhabit the same cultural universe. The notion of orientation to others, in other words, presupposes the cultural capacity of taking of the role or the attitude of the other; and this version of Mead's thesis goes considerably deeper than when the actor simply attempts to view things from the perspective of someone else, in a more or less deliberate manner. What Mead says about children learning the rules of society by playing and participating in games is obviously also relevant in this context.

It seems clear that Mead was trying to solve much larger problems with the notion of taking the role or the attitude of the other than what Weber ever attempted to do with the concept of orientation to others. While Weber was satisfied with having found a concept which allowed

him to sharpen up his own version of sociology, Mead attempted to use the perspective of taking the role or the attitude of the other to explain the very emergence of the self, the way that humans think, the nature of communication, and quite a few other topics as well (e.g. Cook, 1993: 92–6).

From Weber's point of view, Mead's ambitions no doubt appear excessive—but this does not mean that many of Mead's ideas cannot be used to open up new perspectives on Weber's concept of orientation to others. It would appear to be a good working hypothesis, for example, to assume that orientation to others affects the self of the actor; and Mead's ideas on social psychology can more generally be used to counter Weber's negative attitude to the use of psychology in sociology.

Mead was also very interested in perception and symbols; and one wonders how these are related to Weber's notion of orientation to others. Weber avoided these two topics in his theoretical sociology, but not in his sociological studies. Perhaps a close scrutiny of his sociological studies would provide some clues in this context.

In my second example—reference group theory—the term "orientation" is often to be found, although it is not singled out and treated as a special concept of its own. Nonetheless, the idea that an actor orients his or her behavior to some other actor or group is absolutely central to reference group theory.

What is novel about reference group theory is that the actor typically compares or identifies himself or herself to some group, as opposed to looking at oneself and the world through the eyes of somebody else, as in Mead's theory. We have here, in other words, a case of orientation to others which entails a comparison or some form of identification. And secondly, while Mead argues that the individual is formed by the group or society of which he or she is a member, reference group theory suggests that the individual may (also) be influenced by a group of which he or she is *not* a member. A person's orientation, in brief, can be to a membership group as well as a non-membership group; and the latter case may well be the theoretically more interesting one.

While non-membership groups are often spoken of in the social science literature as if they all were of one type, especially Robert K. Merton has emphasized that they are actually quite diverse (see Merton,

[1957] 1968a: 345ff.). There are, for example, non-membership groups which the actor wants to be a member of, as well as non-membership groups which he or she does *not* want to be a member of. In the former case, we may speak of a positive orientation to a non-membership group, and in the latter case of a negative orientation.

There also exist groups that the actor may want to be a member of, but which he or she is not allowed to join; and this means that a reference group can be either open or closed. One can furthermore be oriented positively or negatively to a group, of which one has once been a member. And the remaining members of a group can, in their turn, be positively or negatively oriented to its former members, as such highly charged terms as traitor, renegade and deserter indicate.

Merton also points out that in cases of comparative orientation, the actor has to somehow get to know the norms of the target group, and that the ease or the difficulty with which this can be accomplished needs to be investigated. The degree of social visibility of a group is an important issue from this perspective, and questions of the following type need to be addressed:

> Which processes make for accurate or distorted images of the situation of other individuals and groups (taken as a frame of reference)? Which forms of social organization maximize the probabilities of correct perception of other individuals and groups, and which make for distorted perception? (Merton, [1957] 1968a: 302)

Also some economists have worked with the idea of orientation to others and contributed to its evolution in new directions. This is especially the case with game theory, where the key idea is precisely that an actor takes what another actor does into account in deciding how to act. Much game theory, however, is highly artificial, according to its critics, and cannot be used to capture what goes on in reality.

This type of criticism has especially been directed at game theory by Thomas Schelling, from *The Strategy of Conflict* (1960) and onwards. What Schelling especially directs his criticism at is the so-called two-person zero-sum game. The analysis in this type of game, he says,

operates with far too restrictive assumptions. In particular, Actor A does not need to figure out what Actor B wants, since he or she is supposed to know this in advance. Each player also knows the alternatives available to himself or herself as well as to the other actor (e.g. Luce and Raiffa, [1957] 1985: 57–8). The result is that no orientation to the other needs to take place: "No spark of recognition needs to jump between the two players; no meeting of minds is required; no hints have to be conveyed; no impression, images, or understandings have to be compared" (Schelling, [1960] 1980: 163).

Besides games centered exclusively around conflict, Schelling argues that "a theory of interdependent decision" also should deal with what he calls mixed games and games of coordination. Mixed games—which are probably the most common type of games in empirical reality—have elements of both conflict and coordination. What happens in a pure case of coordination, Schelling specifies, is that Actor A signals something to Actor B in such a way that the latter will expect Actor A to behave in a certain way—and this will also make Actor B behave more predictably from Actor A's perspective. The process of coordination or convergence of expectations is integral to this type of phenomena:

> Players have to understand each other, to discover patterns of individual behavior that make each player's actions predictable to the other … They must communicate by hint and by suggestive behavior. Two vehicles trying to avoid collision, two people dancing together to unfamiliar music, or members of a guerilla force that become separated in combat have to concert their intention in this fashion, as do the applauding members of a concert audience, who must at some point 'agree' on whether to press for an encore or taper off together. (Schelling, [1960] 1980: 85)

In *Micromotives and Macrobehavior* (1978) Schelling makes another contribution to the theory of orientation to others, this time by emphasizing the fact that an actor's behavior may be directly influenced by the exact number of other actors doing something. This last idea has been picked up and further elaborated especially by Mark

Granovetter through the concept of thresholds (e.g. Granovetter, 1978, 1983; Granovetter and Soong, 1988).

A threshold is defined as "the number or proportion of others who must make one decision before a given actor does so"; and it can, according to Granovetter, be used to explain the dynamics of such diverse phenomena voting, migration and the diffusion of innovations (Granovetter, 1978: 1420). Populations with roughly the same average preferences behave, for example, in very different ways, depending on the precise distribution of thresholds among the individuals. If some group lacks people with low thresholds (activists) or with somewhat higher thresholds (excitable people), nothing will happen or the whole process will quickly stall, once it has been set in motion. If these two categories are present, however, the process will start up, accelerate, and perhaps end up in a full-scale riot.

Behavior oriented to coordination with another actor may also involve some kind of signaling, Schelling ([1960] 1980: 83) says; and a whole monograph has been devoted to this phenomenon by economist Michael Spence (1974; for later developments of the theory of signaling, see Gambetta, 2009). According to Spence, two facts characterize signals: they can be altered, and they affect the way that Actor B, who receives the signals, will orient his or her behavior towards Actor A, who has sent the signals. What Spence calls "indices" can, however, *not* be altered, and these refer to such characteristics as sex and race. Under certain circumstances signaling can reach a stable equilibrium, according to a model that Spence specifies. If an employer comes to feel that certain signals that job seekers give off about their productivity, are indeed correct, he or she will be satisfied and stop looking for other signals.

Timur Kuran, another economist, approaches signaling from a very different angle than Spence and in a way that is perhaps more congenial to the sociological version of orientation to others. According to Kuran, an individual may send off false signals, not because he or she wants to reach some specific goal (such as getting hired, in Spence's case), but simply because of social pressure. This phenomenon Kuran terms preference falsification, and he describes it as "misrepresenting one's genuine wants under perceived social pressures" (Kuran, 1995: 3).

Any individual, Kuran argues, has in reality two types of preferences: private preferences and public preferences. In many situations actors prefer to give off false signals, due to social pressure, and in certain cases, as Kuran shows, this can have consequences that the actor deeply regrets. If everybody thinks that everybody else is positive to some legislation, while in reality everyone is against it, the proposal will nonetheless be accepted; and, similarly, if there is no secret ballot, a very unpopular proposal may get passed.

III: CONCLUDING REMARKS

> No one should underestimate the amount of theoretical underpinning by which [Weber's] definition of sociology was sustained. It remains to this day the most consciously chosen, and the most painstakingly derived of any of the multitude of competing definitions which this now prolific intellectual growth has borne.
>
> —Martin Albrow, *Max Weber's Construction of Social Theory*, p. 137

Through the argument in this paper I hope to have established that Weber's concept of orientation plays an important role in his definition of sociology and deserves more attention. I also hope to have shown that other social scientists than Weber have intuitively, and independently of Weber, used ideas that are very similar to the notion of orientation to others, something that again would indicate that we are dealing with an important phenomenon.

As to the research on this topic by other social scientists than Weber, some readers may feel that what Mead *et al.* have to say, rather falls into the well established category of "social interaction" than into the concept of orientation to others. This may well be true to some extent, especially for certain aspects of the work of Mead, but that is not the end of the story. It should be emphasized that Weber *includes* much of what later sociologists term social interaction in his general scheme of sociology. We should recall that Weber starts Chapter 1 of *Economy and Society* with a section on social action and ends with a discussion of organizations—after a section on just social interaction ("social relationships").

More importantly, with the help of the concept of orientation to others Weber can also handle a range of phenomena that the paradigm of social interaction is unable to deal with (e.g. Turner, 1988, Emirbayer, 1998; see also Weber's critique of Simmel's use interaction as the defining feature of sociology in Weber, [1908?] 1912). The Weberian idea of behavior oriented to an order cannot, for example, be explained through the concept of social interaction. The same is true for the key notion in reference group behavior—that an actor orients himself or herself to another group than the one he or she is a member of. In brief, the notion of orientation to others is more subtle than that of interaction.

Many puzzles surrounding the concept of orientation to others no doubt remain to be solved. For one thing, a number of questions relating to Weber's own use of this term need to be answered. What, for example, does it mean to Weber's project of a sociology based on methodological individualism that he suddenly introduces the concept of an *order* and that behavior can be oriented to this order? And why does Weber not explicitly state that an individual's behavior can be oriented to signs and symbols? Does not, for example, the reading of a book—or the reading of this very article—mean that the behavior of the reader is oriented to the ideas of the writer, via the written word? Questions of this type raise interesting and intriguing problems. They should also lead to a realization that we have not yet exhausted Weber's notion of orientation to others and what can be accomplished through his splendid vision of an interpretive sociology.

NOTES

1. I gratefully acknowledge information and/or other help from Mabel Berezin, Hans Henrik Bruun, Laura Ford and Wolfgang Schluchter.
2. There also exist a few other translations into English than the standard one, which comes from the edition of *Economy and Society*, edited by Guenther Roth and Claus Wittich (Weber, [1922] 1978). Before citing these other translations, I will however cite the full comment on the word "orientieren" that one can find in Hans Henrik Bruun's translation of the *Wissenschaftslehre* and related texts into English:

> *orientieren*: The obvious translation is 'orient'. But, in *Categories* (RS: Weber, [1913] 2012), 'orientieren an' can be a problem. In many cases, the

> 'orientation' means that the acting person somehow relates his action to something, without specifying whether that 'orientation' represents conformity or the reverse. (Indeed, Weber specifically discusses the case of criminals 'orienting' themselves towards an established order, not by conforming, but by trying to conceal their crime). In these cases, 'orientieren an' has been translated as 'orient towards'. But there are instances where the orientation is very definitely meant to represent conformity. In those cases, 'orientieren an' has been translated as 'orient in accordance with' or 'orient according to'. (Weber, 2012: 491).

Talcott Parsons translated the key sentence in Weber's definition of sociology as follows: "Action is social in so far as, by virtue of the subjective meaning attached to it by the acting individual (or individuals), it takes account of the behaviour of others and is thereby oriented in its course" (Weber, 1947: 88). Parsons also attached a long footnote that singles out and discusses a number of terms—but not "orientation to others" ("In this series of definitions Weber employs several important terms which need discussion. In addition to *Verstehen*, which has already been commented upon, there are four important ones: *Deuten*, *Sinn*, *Handel*n, and *Verhalten*" (Weber, 1947: 88, note 3). The translation by H. P. Secher reads as follows: "The term 'social behavior' will be reserved for activities whose intent is related by the individuals involved to the conduct of others and is oriented accordingly" (Weber, 1962: 27). Keith Tribe's translation: "Such behaviour is 'social' action where the meaning intended by actor or actors is related to the behaviour of *others*, and conduct so oriented" (Weber, 2004: 312).

3. Hopefully there will be some comments in the *Gesamtausgabe* of Weber's works. Ch. 1 *in Economy and Society* is scheduled to appear in Band 23, which had not been published when this paper was written (January 2013). The title of Band I/23 is *Wirtschaft und Gesellschaft. Soziologie Unvollendet 1919–1920*. According to Wolfgang Schluchter, Band I/23 will appear in the summer or fall of 2013.
4. For Parsons' translation of Ch. 1 in *Economy and Society*, see note 2.
5. See e.g. Girndt (1967: 44–6). In Winkelmann's well-known volume with comments on *Economy and Society* there is no reference to the part of Weber's definition of sociology that deals with orientation to others (Winkelmann, 1976: 1). The literature on social action similarly avoids discussing the element of orientation to others in Weber's definition (see e.g. Campbell, 1996).
6. "A proto-concept", to cite Merton's definition, "is an early rudimentary, particularized, and largely unexplicated idea …; a concept [on the other hand] is a general idea which once having been defined, tagged, substantially generalized, and explicated can effectively guide inquiry into seemingly diverse phenomena" (Merton, 1984: 267).
7. Kant mentions several uses of the term orientation in his essay "What is Orientation in Thinking?" There is first of all "the proper sense of the

word", which Kant also calls "the geographical concept of the process of orientation". He describes this as follows: "to *orient* oneself, in the proper sense of the word, means to use a given direction—and we divide the horizon into four of these—in order to find the others, and in particular that of *sunrise*" (Kant, 1991: 238–39). One may also "extend" the geographical concept for use in mathematics, Kant says; and this "*mathematica*l sense [of the word orientation]" constitutes its second meaning in his work. "Finally", Kant writes, "I can extend this concept even further" and this would be to "to orient oneself … in thought, i.e. *logically*" (Kant, 1991: 239).

This last meaning is clearly the one Kant was most interested in and he also describes it as follows: "If the concept of orientation is extended and defined more precisely, it may help us to cast light on the various ways in which the maxim of healthy reason is applied to the cognition of supra-sensory objects" (Kant, 1991: 238).

In the entry for "orientation" in *A Kant Dictionary*, Howard Caygill writes that "the notion of orientation is used [by Kant] in theoretical philosophy to pass judgment on the speculative use of reason with respect to the supersensible" (Caygill, 1995: 311).

Finally, the argument in an interesting article by Stuart Dalton on orientation suggests a possible link between Kant and Weber, namely that subjectivity can be inner-oriented or other-directed (to use the terminology of David Riesman). Or in Dalton's own terminology, one may, with the help of Levinas and Kant, view "subjectivity as orientational rather than foundational in nature" (Dalton, 1999: 433).

8. In the definition in Chapter 1 in *Economy and Society* Weber uses "*orientiert*" as well as "*bezogen*": "Action is 'social' insofar as its subjective meaning takes account of (*bezogen*) the behavior of others and is thereby oriented (*orientiert*) in its course" (Weber, [1922] 1978: 4; for the German original, see Weber, [1922] 1972: 1). "Related to" in the definition from *Logos* is a translation of "*bezogen*"; see Weber, [1913] 2012: 274; for the German original, see Weber, [1913] 1922: 429. The term "oriented to" is also used a number of times in the *Logos* article.
9. Value-rational action is per definition directed at a value in a specific way, and it becomes social, as I interpret Weber, by simultaneously being oriented to the behavior of some actor(s). An example of a *social* value-rational action would be when someone who thinks it is absolutely right/wrong that abortion is legal, says so at a public meeting of opponents.
10. There exist two kinds of economically oriented actions in Weber's economic sociology in *Economy and Society*: economic action which makes use of violence and non-economic actions which also takes economic considerations into account.
11. Weber does not use the concept of society in his theoretical sociology. See e.g. Frisby Sayer (1986: 67–72).

12. Other examples exist as well. That the concept of "field", as used in the work of Pierre Bourdieu, includes the notion of orientation to others is clear (see e.g. Bourdieu and Wacquant, 1992: 94–115). Also the new institutionalists in organization theory argue that all organizations in an organizational field are aware of each other and that this awareness is constitutive for the very existence of the field (e.g. DiMaggio and Powell, 1991: 64–5; cf. DiMaggio, 1986). For an argument that markets come into being through the actors being oriented to one another through terms of trade-schedules, see White (1981); and for the idea that corporations diffuse "conceptions of control" by being oriented to each other, see Fligstein (1990).

REFERENCES

Albrow, Martin (1990) *Max Weber's Construction of Social Theory*. London: Macmillan.

Aron, Raymond ([1967] 1970) *Main Currents of Sociological Thought II. Durkheim, Pareto, Weber*. trans. Richard Howard and Helen Weaver. New York: Anchor Books.

Bourdieu, Pierre and Loïc Wacquant (1992) *An Invitation to Reflexive Sociology*. Chicago: University of Chicago Press.

Campbell, Colin (1996) *The Myth of Social Action*. Cambridge: Cambridge University Press.

Caygill, Howard (1995) "Orientation", in Howard Caygill, *A Kant Dictionary*. Oxford: Blackwell, p. 131.

Cherkaoui, Mohamed (2007) *Good Intentions: Max Weber and the Paradox of Unintended Consequences*. Oxford: The Bardwell Press.

Cook, Gary A (1993) *George Herbert Mead: The Making of a Social Pragmatist*. Urbana: University of Illinois Press.

Dalton, Trevor (1999) "Subjectivity and Orientation in Levinas and Kant", *Continental Philosophy Review* 32: 433–49.

DiMaggio, Paul (1986) "Structural Analysis of Organizational Fields: A Blockmodel Approach", *Research in Organizational Behavior*, 8: 335–70.

DiMaggio, Paul and Walter W. Powell (1991) "The iron cage revisited: Institutional isomorphism and collective rationality in organizational fields", in DiMaggio and Powell (eds.), *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press, pp. 62–82.

Emirbayer, Mustafa (1997) "Manifesto for a relational sociology", *American Journal of Sociology* 103: 281–317.

Fligstein, Neil (1990) *The Transformation of Corporate Control*. Cambridge: Harvard University Press.

Freund, Julien ([1966] 1972) *The Sociology of Max Weber*. Trans. Mary Ilford. Harmondsworth: Penguin Books.

Frisby, David and Derek Sayer (1986) *Society*. London: Tavistock Publications.

Gambetta, Diego (2009) "Signaling", in Peter Hedström and Peter Beramn (eds), *The Oxford Handbook in Analytical Sociology*. New York: Oxford University Press, pp. 168–94.

Girndt, Helmut (1967) *Das soziale Handeln als Grundkategorie erfahrungswissenschaftlicher Soziologie*. Tübingen: J. C. B. Mohr.

Granovetter, Mark (1978) "Threshold models of collective behavior", *American Journal of Sociology*, 83: 1420–43.

——— (1983) "Threshold models of diffusion and collective behavior", *Journal of Mathematical Sociology*, 9: 165–79.

Granovetter, Mark and Roland Soong (1988) "Threshold models of diversity: Chinese restaurants, residential segregation, and the spiral of silence", *Sociological Methodology*, 18: 69–104.

Joas, Hans (1985) *G. H. Mead: A Contemporary Re-Examination of His Thought*. Trans. Raymond Meyer. Cambridge: Polity Press.

Kant, Immanuel ([1786] 1991) "What is Orientation in Thinking?" in Immanuel Kant, *Political Writings*. Ed. H. S. Reiss. Cambridge: Cambridge University Press, pp. 237–49.

Käsler, Dirk (1988) *Max Weber: An Introduction to His Life and Work*. Trans. Philippa Hurd. Cambridge: Polity Press.

Kuran, Timur (1995) *Private Truths, Public Lies: The Social Consequences of Preference Falsification*. Cambridge: Harvard University Press.

Mead, George H. ([1934] 1962) *Mind, Self, and Society: From the Standpoint of A Social Behaviorist*. Chicago: University of Chicago Press. Based on a course given in 1927.

Merton, Robert K. ([1948–1949] 1968) "Patterns of influence: Local and cosmopolitan influences", in *Social Theory and Social Structure*. 1968 enlarged edn. New York: The Free Press, pp. 441–74.

——— ([1957] 1968a) "Continuities in the theory of reference groups and social structure", in *Social Theory and Social Structure*. 1968 enlarged edn. New York: The Free Press, pp. 335–440.

——— ([1957] 1968b) "(With Alice Rossi) Contributions to the theory of reference group behavior", in *Social Theory and Social Structure*. 1968 enlarged edn. New York: The Free Press, pp. 279–5.

——— (1984) "Socially expected durations: A case study of concept formation in sociology", in W. W. Powell and R. Robbins (eds), *Conflict and Consensus*. New York: The Free Press, pp. 262–83.

Parsons, Talcott ([1937] 1967) *The Structure of Social Action*. Vol. II. New York: The Free Press.

Schelling, Thomas ([1960] 1980) *The Strategy of Conflict*. Cambridge: Harvard University Press.

——— (1978) *Micromotives and Macrobehavior*. New York: W. W. Norton & Company.

Schelting, Alexander von (1934) *Max Webers Wissenschaftslehre*. Tübingen: J. C. B. Mohr.

Schluchter, Wolfgang (1998) "Max Webers Beitrag zum *'Grundriss der Sozialökonomik'*", *Kölner Zeitschrift für Soziologie und Sozialpsychologie*, 50: 327–43.

Schutz, Alfred ([1932] 1967) *The Phenomenology of the Social World*. Trans. George Walsh and Frederick Lehnert. London: Heinemann Educational Books.

Spence, Michael (1974) *Market Signaling: Informational Transfer in Hiring and Related Screening Processes*. Cambridge: Harvard University Press.

Swedberg, Richard (2005) "Orientation to others", in *The Max Weber Dictionary*. Stanford: Stanford University Press, pp. 190–91.

Turner, Jonathan H (1988) *A Theory of Social Interaction*. Cambridge: Polity Press.

Weber, Max ([1908?] 2012) "Georg Simmel as a sociologist and theoretician of the money economy (trans. Hans Henrik Bruun)", in Max Weber, *Collected Methodological Writings*. Trans. and ed. Hans Henrik Bruun and Sam Whimster. London: Routledge, pp. 418–21.

——— ([1913] 1922) "Ueber einige Kategorien der verstehenden Soziologie", in *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen: J. C. B. Mohr, pp. 403–50.

——— ([1913] 1981) "Some Categories of Interpretive Sociology (trans. Edith Graber)", *Sociological Quarterly*, 22: 151–80.

——— ([1913] 2012) "On Some Categories of Interpretive Sociology (trans. Hans Henrik Bruun)", in Max Weber, Max, *Collected Methodological Writings*. Trans. and ed. Hans Henrik Bruun and Sam Whimster. London: Routledge, pp. 273–301.

——— ([1922] 1972) *Wirtschaft und Gesellschaft. Grundriss der verstehenden Soziologie*. 5th rev. edn. Johannes Winkelmann. Tübingen: J. C. B. Mohr.

——— ([1922] 1978) *Economy and Society: An Outline of Interpretive Sociology*. Guenther Roth and Claus Wittich (eds). Trans. Ephraim Fishoff *et al*. Berkeley: University of California Press.

——— (1947) The *Theory of Social and Economic Organization*. Ed. and trans. Talcott Parsons. New York: The Free Press.

——— (1962*) Basic Concepts in Sociology*. Trans. H. P. Secher. New York: The Citadel Press.

——— (2004) "Basic sociological concepts" (trans. Keith Tribe), in Sam Whimster (ed.), *The Essential Weber: A Reader*. London: Routledge, pp. 311–58.

——— (2012) *Collected Methodological Writings*. Trans. and ed. Hans Henrik Bruun and Sam Whimster. London: Routledge.

White, Harrison C (1981) "Where do markets come from?" *American Journal of Sociology*, 87: 517–47.

Winkelmann, Johannes (1976) *Wirtschaft und Gesellschaft. Erläuterungsband*. Tübingen: J. C. B. Mohr.

PART TWO

Epistemology of the Social Sciences

5

GENERATIVE MECHANISMS AND DECREASING ABSTRACTION[1]

Jack Birner

University of Trento; University College Maastricht;
Konrad Lorenz Institute for Evolution and Cognition Research

> One of the questions I pose is that of understanding how the classical sociologists and we ourselves construct explanatory schemas that aim to integrate the macrosociological and the microsociological levels and how to construct a research strategy based around the principles of methodological individualism for understanding macro phenomena. The transitions between these levels operate through the means of a construction and modelling of generative mechanisms. (Cherkaoui, 2005: 1)

INTRODUCTION

In *Invisible Codes: Essays on Generative Mechanisms* Mohamed Cherkaoui addresses two major problems. One is the integration of empirical research and theoretical analysis in sociology, which, he says, has not yet been solved.[2] The second problem is the analysis of generative mechanisms and the methodology that is suitable for studying

them. Generative mechanisms are processes that transform individual actions into collective phenomena. An important subset of these are social amplifiers: mechanisms that transform individual human actions into non-linear intended and unintended collective effects. In my opinion, explaining social amplifiers, or invisible-hand mechanisms, is the most important task of the social sciences. The capacity to explain the causal mechanisms that underlie them is what makes social science more than the sum of psychology and arithmetic;[3] it makes it a scientific discipline in its own right.

Social amplifiers, however, remain under-analysed because of a neglect of either the role of individual agents or the social structure that binds them together. In economics, the role of social structures, or networks, is still insufficiently understood, as is the role of individuals in sociology. As to the causal mechanisms involved in social amplifiers, neoclassical economics locates them in the behaviour of individuals.[4] The network sociologist Ron Burt, on the contrary, attributes causality exclusively to the structure of social interaction.[5] Modelling the structure in which individual agents interact is a necessary condition for the analysis of social amplifiers: without interaction there can be no amplification of the effects of individual actions. It is likewise with explanations that refer to structural causes; they need individual agency to set them in motion. Social scientists seem to find it difficult, though, not to succumb either to micro-reductionism, as research in experimental economics, cognitive science and neuro-economics illustrates, or to its opposite, which is the prevalent approach in much of the research in sociological network theory. In addition to being radically reductionist, these approaches are causally incomplete.[6]

This brief discussion allows us to reformulate Cherkaoui's second problem as the question how to integrate the causal influences of individual agency and social structure in a single explanatory framework. I hope to show that both of Cherkaoui's problems may be solved by adding an instrument to his methodological toolkit, which, like that kit, has firm roots in the 1930s. For that purpose I will discuss the philosophy of social science of Karl Popper and Friedrich von Hayek. This will also shed light on a couple of links in the history of the methodology of social science that have remained a bit in the shade. The conclusion of

the following paragraphs will be that the method of decreasing abstraction has an important role to play in sociology. More in particular, it is instrumental for solving Cherkaoui's two problems.

THEORY AND EMPIRICAL EVIDENCE: POPPER'S METHODOLOGY OF SOCIAL SCIENCE

In *Logik der Forschung* (1934), Karl Popper had shown how empirical data and theories are related in natural science. Subsequently, he tried to extend his philosophy to the social sciences. He was influenced in this by Friedrich von Hayek.[7] After reading Popper's *Logik der Forschung*, Hayek realized that the methodology he had adopted from Ludwig von Mises failed to give a satisfactory account of the relationship between empirical and theoretical economics. He put this problem on his agenda. The result was "Economics and Knowledge," published in 1937.

Popper and Hayek had met the year before. Popper, who was much less familiar with social than with natural science, was happy to let himself be guided by Hayek, who, after all, was an economist. This influence, however, was not entirely positive. For reasons I will explain, Popper, who had given a very clear explication of the relationship between theory and empirical observation in the natural sciences, failed to do the same for the social sciences. This is somewhat paradoxical as his work contains the elements for solving the problem in a satisfactory manner. That has been demonstrated by Noretta Koertge. These issues will be addressed in more detail below.

In Popper's and Hayek's sociological theories the idea that individual actions typically have unintended consequences plays a central part. Unintended consequences are closely related to social amplifiers[8] or invisible-hand mechanisms. I will make a comparison between Cherkaoui's methodology of generative mechanisms on one hand and Karl Popper's sociology and its methodology—to which Cherkaoui refers—and Hayek's method on the other. The links between Popper and Hayek and more recent methodological developments show the way to the correct framework for the explanation of invisible-hand processes.

It is useful to distinguish three aspects of social amplifiers: their ontology (*i.e.*, the existence of relevant causal mechanisms), the way to

discover and study them (*i.e.*, the heuristics, epistemology and methodology associated with their analysis), and the way in which models and theories about them may be formalized.[9] An important function of methodology, possibly reinforced by the choice of formalism, is to serve as a check on ontological proliferation. This is clearly one of the purposes of the core of Popper's methodology of social science, the rationality principle.[10] The following bit of intellectual historiography will help to make this clear.

Popper and Hayek first met in 1936. In one of the seminars that Hayek organized at the London School of Economics, Popper presented some of the arguments that he would later publish in *The Poverty of Historicism*. Their encounter was the beginning of an exchange of ideas that lasted for the rest of their lives and influenced the thought of both.[11] For one thing, it led Hayek to write "Economics and Knowledge," in which he tried to reconcile Ludwig von Mises' idea that economics was "a priori" with Popper's argument that in order to be scientific, theories must have empirical content.[12] As to Popper, his contact with Hayek was a welcome opportunity for learning more about economics.[13] Hayek's habit of systematically accompanying his scientific endeavours with an exposition of the methodology that is appropriate for them was a further stimulus to Popper's curiosity about the methodology of the social sciences.

So when Popper started elaborating his methodology of social science, Hayek's methodology of economics was an important point of reference. That is confirmed in the opening statements of Popper's "Models, Instruments and Truth. The status of the rationality principle in the social sciences", which is based on a lecture he gave in 1963: "My views on the methodology of the social sciences are the result of my admiration for economic theory: I began to develop them, some twenty-five years ago, by trying to generalize the method of theoretical economics." (Popper, 1963: 181). In a footnote he adds:

> I was particularly impressed by Hayek's formulation that economics is the 'logic of choice' (…). This led me to my formulation of the 'logic of the situation' (cf. My *Poverty of Historicism*, p. 149). This seemed to me to embrace, for example, the logic of

> choice and the logic of historical problem situations. (The origin of this idea may explain why I rarely stressed the fact that I did not look at the logic of the situation as a deterministic theory: I had in mind the logic of situational choices.) (*ibid*.)

Hayek's influence also shows in the distinction Popper makes between the rationality principle and the description of an individual's social situation. This will be explained in what follows.

Popper observes that theories may either explain singular events or types of events. In the former case, universal theories and relevant singular initial conditions are needed; in the latter, we need models, which include the description of typical initial conditions.[14] Though the use of models is characteristic of social science, they are also used in natural science: "It seems that, as a rule, we first operate with models, and that the models, *together with a rudimentary working mechanism,* may solve a number of problems of the second kind—that is, explain some typical events." (Popper, 1963: 165). But, he says, in social science the use of models is more important. "This is connected with the fact that in the social sciences there is, in Hayek's terminology, less "explanation in detail" and more "explanation in principle" than in the physical sciences." (Popper, 1963: 166). Thus, the use of models constitutes a difference in degree between natural and social science.

But there is also a difference in kind: "The fundamental problem of both the theoretical and the historical social sciences is *to explain and understand events in terms of human actions and social situations.* The key term here is "social situation"". (Popper, 1963: 166). The description of social situations is part of the initial conditions. It makes an appeal to psychological facts superfluous:

> I propose to treat [an individual's] aims and (…) knowledge not as psychological facts, to be ascertained by psychological methods, but as *elements of the objective social situation*. (Popper, 1963: 167)

To which he adds:

> We shall be concerned only with information or knowledge (…) that is relevant to the situation. (Popper, 1963: 168)

Not only is there no need for psychological laws; invoking them would be a mistake. First, because "we have already *replaced* [the individual's] concrete, conscious or unconscious, psychological experiences by some abstract and typical situational elements, such as those we dubbed "aims" and "knowledge"." (Popper, 1963: 169). And second, because

> it is the central point of situational analysis that we need, in order to "animate" it, no more than the assumption that the various persons or agents involved act *adequately*, or *appropriately*—that is to say, in accordance with the situation. We must remember, of course, that the situation, as I use the term, already contains all the relevant aims and all the available knowledge, especially of the various possibilities for realizing these aims. […]
>
> The adoption of the rationality principle can therefore be regarded as a by-product of a methodological postulate. It does not play the role of an empirical explanatory theory, of a testable hypothesis. For in this field, the empirical explanatory theories or hypotheses are, rather, our various models, our various situational analyses. It is these which may be empirically more or less adequate—which may be discussed and criticized, and whose adequacy may sometimes even be tested, thereby enabling us, if they fail the test, to learn from our mistakes. (Popper, 1963: 169–70)

Popper shares these ideas about the function of hypotheses about individuals and their situation in explaining collective phenomena with Hayek. Their position is known as anti-psychologism. I will return to it below.

In *Logik der Forschung* Popper had emphasized the importance of severe tests. It therefore comes as a bit of a surprise that he relaxes this requirement when it comes to the rationality principle: "Tests of a model, it has to be admitted, are not easily obtainable and are usually

not very clear-cut." (Popper, 1963: 170). He even goes so far as to extend this to the testability of theories in natural science: "this difficulty arises even in the physical sciences." (*ibid.).* Its non-testability is not the only reason why Popper's rationality principle has been criticized.[15] Critics have observed that Popper's account of the rationality principle is marred by inconsistencies: the rationality principle is proposed as a methodological principle, but it is also empirical.[16] It has little empirical content, is false, yet a good approximation to the truth, but it does not play the role of a testable hypothesis. And despite its falsehood, it must not be rejected or replaced by a hypothesis that is closer to the truth, which is in blatant contradiction with Popper's falsificationism. Where do these confusions come from?

POPPER AND HAYEK

Their origins lie in Hayek's work. What Popper writes in the passages just quoted reflects the distinction Hayek makes, in "Economics and Knowledge," between a tautological and an empirical part of economic theory. So, in order to understand what Popper tries to do—in response to what he had learnt from Hayek—let us return to what Hayek thought he had learnt from Popper. Mises thinks that economics is "a priori." He appears to mean by this that economic theory is a priori true or tautological. This is inconsistent with the requirement, so forcefully exposed by Popper in *Logik der Forschung*, that scientific theories have empirical content. Hayek tries to reconcile Mises with Popper by recognizing the existence of the tautological pure logic of choice but also of a separate part of economic theories that contains their empirical content. It consists of the hypotheses about the way individuals acquire information.[17] So Hayek does not eliminate tautologies from economic theory; he supplements them with empirical statements. In the perspective of the *Logik der Forschung* this solution is not satisfactory,[18] but I will let that rest for now. What is of more interests for my argument here is that it is difficult *not* to recognize Hayek's distinction behind the one Popper makes between the verisimilar-yet-false-and-untestable-yet-not-to-be-abandoned rationality principle (to mention only a few of the confusions Popper makes), on one hand, and the description of the situation on the other.

So, what we see here is the following extraordinary and curious episode. Hayek writes "Economics and Knowledge" in an attempt to integrate Popper's concept of empirical content within his own methodology. Then Popper takes this article as the point of departure for an elaboration of his—Popper's—methodology of social science. Hayek's translation of Popper's philosophy of science into a Misesian methodological framework fails to make clear what exactly the relationship is between the a-priori and the empirical parts of economic theory. Popper's retranslation of this translation of Hayek's back into his own falsificationist philosophy of science is, however, even less satisfactory. It gives rise to confusions and inconsistencies. Nevertheless, Hayek's attempt to add hypotheses with empirical content to the tautological pure logic of choice is a step forward. It is a move away from Mises' apriorism towards a methodology that is capable of accounting for the requirement that economic theory have empirical content, and hence—*pace* Popper—be scientific. For Popper, on the contrary, the attempt to integrate Hayek's methodology into his own philosophy of science is a step backward. It is a move away from the clear requirement that theories be falsifiable, towards what may easily be interpreted as a variant of instrumentalism, a methodology Popper rejects.

Popper was well aware of the danger of importing instrumentalism together with other ideas that he took over from Hayek (or for which he was were inspired by him). That is shown by the title, "Models, Instruments and Truth," and by the fact that he dedicates four sections (8–11) to rejecting instrumentalism. He starts section 8 by observing: "there is a group of philosophers who may (somewhat precipitately) be pleased about what I have been saying: the pragmatists or instrumentalists." (Popper, 1963: 171). Hayek is not mentioned, despite the fact that he openly embraces an instrumentalist methodology. He presents it at length in *The Counter Revolution of Science*. Another place where he does so is "The Facts of the Social Sciences." There he writes that

> the theories of the social sciences do not consist of "laws" in the sense of empirical rules about the behavior of objects definable in physical terms. All that the theory of the social sciences attempts is to provide a technique of reasoning which assists us

> in connecting individual facts, but which, like logic or mathematics, is not about the facts. It can, therefore (...) never be verified or falsified by reference to facts. All that we can and must verify is the presence of our assumptions in the particular case. (...) In this connection a genuine "question of fact" arises - though one it will often not be possible to answer with the same certainty as in the case of the natural sciences. But the theory itself, the mental scheme for the interpretation, can never be "verified" but only tested for its consistency. It may be irrelevant because the conditions to which it refers never occur; or it may prove inadequate because it does not take account of a sufficient number of conditions. But it can no more be disproved by facts than can logic or mathematics. (Hayek, 1943: 73)[19]

Popper concedes that instrumentalism is an internally consistent methodology. The reason why he rejects it is that it cannot account for the explanatory function of science. In its place he proposes the idea of a series of successive theories each of which is a closer approximation to the truth. In section 10 of "Models" he uses the same example as in "The Aim of Science" of 1957. That article analyzes idealizations and approximation to the truth in natural science. But in "Models" Popper does not elaborate this. For lack of further argument his statement that his methodology of social science is not instrumentalist fails to convince, as do his insistence that the rationality principle, though false but a good approximation to the truth, should not be substituted by a more verisimilar hypothesis and his emphasis on its the unifying function. The conclusion is warranted that Popper, against his express will, is stuck with at least some of the instrumentalist baggage that comes with the rest of Hayek's methodology.

As if this were not bad enough, Hayek's negative influence does not stop here. We have seen that Popper liked Hayek's idea of the logic of choice as the core of economic theory. At the same time, he rejects the idea that situational analysis is deterministic.[20] But as Latsis has shown,[21] neoclassical microeconomics, which is built on the theory of rational choice (the "pure logic of choice") together with its other standard assumptions leaves individuals without any choice: its models

indicate a unique "single exit" out of what is presented as a situation in which individuals have a choice. So, a further consequence of following Hayek's methodology is that Popper is stranded with yet another doctrine he rejects, determinism. These are perhaps the two most disastrous consequences of Popper's deriving his inspiration from Hayek.

There is, however, an alternative and more benign way of looking at Popper's conundrum. It is that, given that much of his knowledge of economics and its methodology had come from Hayek, Popper made the most of an unclear situation. For the second time,[22] Popper tried to draw the consequences from the fact that theories—in this case in social science—are idealizations. Compare, for instance, the passage in "Models" where he writes that the theory of perfect competition in economics and the "pure theory of monopoly or duopoly is nothing but the situational logic of certain idealized social situations." (Popper, 1963: 170). As I have observed above, in "The Aim of Science" of 1957 he had addressed the problem of idealizations in natural science. In that article he proposes his "correspondence principle", which is a sound basis for a theory of idealizations.[23] The *Poverty of Historicism*, published as a book in the same year, contains a rather strong hint at the idealizing character of the "zero method":

> I refer to the possibility of adopting, in the social sciences, what may be called the method of logical or rational reconstruction, or perhaps the 'zero method' By this I mean the method of constructing a model on the assumption of complete rationality (and perhaps also on the assumption of complete information) on the part of the individuals concerned, and of *estimating the deviation of the actual behavior,* using the latter as a kind of zero coordinate. (Popper, 1957: 141, my italics).

In a footnote he adds: "perhaps, … the use of rational or 'logical' models in the social sciences, or of the 'zero method,' has some vague parallel in the natural sciences …" (*ibid.*). Yet despite all these "markers" of idealizations, he never pursues these parallels nor does he elaborate a methodology of idealizations. It remains a neglected topic in Popper's philosophy of science.

SAVING POPPER FROM HIMSELF—BY HIS OWN MEANS

In what is in fact an application of the methodology of idealizations to social science (though it is not explicitly presented as such), Noretta Koertge proposes a solution to Popper's muddled discussion of the rationality principle. The crux of her proposal is to interpret the idea that individuals act rationally as an idealizing empirical hypothesis that may be factualized.[24] She shows that, by applying a variant of Popper's correspondence principle (to which she does not refer either), the rationality principle may be given empirical content through a series of factualizations. This is done by supplementing the hypothesis that individuals act rationally by a number of auxiliary hypotheses, which act as initial conditions.[25] The following explanatory scheme shows how.

1. Description of problem situation:	Individual A thought he was in problem situation of type C.
2. Dispositional law:	For all such problem situations A would use appraisal rule R.
3. Analysis of the situation:	The result of A appraising C using R is x.
4. Description of agent's competence:	A did not make a mistake in applying R to C.
5. Explanandum-1:	Therefore A concluded that x was the rational thing to do.
6. Rational appraisal principle:	All agents appraise their situations in a rational manner.
7. Rational action principle:	People always act on the outcome of their rational appraisals.
8. Explanandum-2:	Therefore A did x.

Koertge argues that her scheme increases the empirical content of the rationality principle because the conjunction of the rationality principle with the auxiliary theories "rules out a large number of possible worlds."[26] For this to be the case it is necessary that independent evidence can be produced for each of the initial conditions. If not, the explanation

would be *ad hoc* and hence content-decreasing. Falsifications should be attributable to one or more of the auxiliary assumptions by explaining *why* they failed to be true. This requires additional explanatory schemes that explain and predict the auxiliary assumptions with the help of further explanatory laws.

Koertge's proposal shows striking similarities with two features of Hayek's work. This is surprising as there is no evidence that Koertge was familiar with it. She may have inherited these aspects of Hayek's methodology from Popper's rationality principle. One similarity concerns Hayek's idea that the empirical content of economics is to be sought in the hypotheses about how individuals acquire information.[27] Koertge's scheme leaves ample space for—even invites—additional hypotheses about the way in which individuals perceive their environment and may correct mistaken perceptions.[28] In addition, her elaboration of Popper's rationality principle is an application of the method of decreasing abstraction that has a central role in Hayek's analytical work.[29] As the remainder of this chapter is dedicated to this method, let me first illustrate how he uses it. For that purpose, I will discuss Hayek's "Das intertemporale Gleichgewichtssystem der Preise und die Bewegungen des 'Geldwertes'" of 1928 and "Economics and Knowledge" of 1937. In the above, I have repeatedly referred to the latter for its methodological contents; in what follows I shall concentrate on the economic theory it contains. Readers who find the example too long or too technical may skip it and go directly to the summary in the last paragraph of the following section.

THE REPRESENTATIVE INDIVIDUAL

In order to understand how the two articles are related, a few words on Hayek's problem situation are in order. From the first half of the 1920s, the main item on Hayek's research agenda was the explanation of business cycles, the alternation of periods of accelerated economic growth with episodes of stagnation or even negative growth. From Ludwig von Mises and Knut Wicksell he takes over the idea that monetary factors such as the interest rate banks charge for loans, the difference between these and the real rate of return on investment (or productivity), and

variations in the quantity of money play a central role. But differently from Mises, Wicksell and some other contemporary economists (notably Wesley Mitchell), Hayek rejects the general price level as an explanatory factor. In addition to this self-imposed theoretical constraint, he adopts another: he wants the explanation to remain within the framework of the general equilibrium theory that had been developed by Léon Walras and Vilfredo Pareto. That theory describes how all markets in an economy are related to one another. The link lies in the mechanism that coordinates the choices of individuals on the basis of price signals. This methodological choice of Hayek's creates two further problems for him. First, he has to find a way to make general *equilibrium* theory applicable to the *disequilibrium* phenomenon of the business cycle. Second, general equilibrium theory is a *static* (or better: timeless) theory in the sense that it describes the relationships between markets at one moment in time. What Hayek wants to explain, however, is disequilibrium growth, a *dynamic* phenomenon that involves changes over a period of time. These are the preliminary problems he had to solve before giving an explanation of the business cycle.

Hayek solves them by introducing a number of methodological devices. He simplifies the model by abandoning the objective of including the relationships between all markets in favour of a model that describes the relationships between the demand for and supply of investments and savings and the demand for and supply of one (presumably composite) consumer good. At the same time, he generalizes general equilibrium theory into a theory that describes the relationships between these two markets (for savings-investments and the consumer good) through time. The result is the first intertemporal general equilibrium theory in the history of economics. Time is introduced by a very ingenious device: Hayek chooses as the basic unit of analysis the individual plan. The concept of a plan logically implies time (timeless planning is a contradiction in terms). Hayek introduces an additional idealization, one that plays a pivotal role: the representative individual. It is important for two reasons. First, it helps him achieve the main purpose of the article of 1928, to explain how money and prices influence the relationships between supply and demand over time. Second, it serves as a link with a second problem he wants to address: how to model the

coordination between individuals, as a preliminary to the analysis of the *lack* of coordination that causes business cycles. The link is activated, so to speak, by abandoning the idealization of the representative individual in favour of a more factualized model with many individuals. I will return to this below.

Let us go back to the 1928 article. The elements described above lead to a definition of equilibrium as the situation in which the representative individual receives no information that is inconsistent with the data on which he had based his plan. In other words, the individual is in equilibrium when he has correctly foreseen economic reality, or, what is the same, if reality behaves according to his expectations. If this is the case, there is no need for the individual to adapt his plan. This concept of equilibrium is explicitly introduced as a theoretical fiction, one that is not intended to describe reality. It serves the methodological purpose of acting as a point of reference for more realistic models of an economy. Another methodological device that Hayek introduces is the distinction between a hypothetical barter economy, where relative prices are in terms of exchange relationships between goods,[30] and an economy with money and thus with prices quoted in dollars or euros or whatever the currency is.

This extensive toolbox of methodological devices enables Hayek to answer the question about the causes of business cycles. Very briefly,[31] his solution is that in a partial-reserve banking system (which is the one in vigour in all modern economies) the money (or credit) multiplier has an autonomous influence on money prices in addition to the relative real prices (which would be the only ones to exist in a barter economy). It acts both as an exogenous source of instability and as an amplifier of real shocks. As a consequence, money prices do not reflect real scarcities. Since the individual designs his plan on the basis of the only prices he can perceive, money prices, his decisions to consume, save and hence consume in the future (in an intertemporal model, savings equal deferred consumption), and invest do not reflect real scarcities. The expectations on which the individual plan is based are falsified for reasons that lie in the monetary sector of the economy. This falsification is tantamount to a disequilibrium of the individual decision maker, who therefore adapts his plan. The result is that the quantity of the consumer

good that will be produced in the future, which is a product of the investments that have been made earlier, and which in their turn are based on the individual's past expectations, will not match the demand for the consumer good at the same moment in the future. This causes an intertemporal misallocation of investments or a coordination failure between supply and demand over time. Business cycles are the consequence of failures of intertemporal coordination between demand and supply of goods. These coordination failures are caused or amplified by the monetary system.

In 1928 Hayek indicates that this leaves another problem to be solved: how to explain the coordination failures between the numerous individuals of which an economy consists. That is the main economic problem he addresses in 1937. It is a step in the direction of a more realistic model of an economy, or a factualization of his more idealized model of 1928. In order to address the coordination problem between different individual plans, Hayek must, of course, abandon the idealizing assumption of the representative individual. The result is a model in which the price signals generated by the interaction between individuals on markets are the informational inputs on which all individuals calibrate their plans and hence their market behaviour. One consequence (or corollary or added theoretical value) of this factualization is that it indicates that the position of a particular individual occupies in the interaction and communication structure of the entire economy co-determines the information he has access to. The image Hayek has in mind in his 1937 is that of a network model *avant-la-lettre* of an exchange economy.[32] All of this enables him to define the equilibrium of the economy as the mutual compatibility between all individual plans. Again, he makes it clear that this is a hypothetical concept that does not pretend to describe an observable state of affairs in a real economic system.

Let me summarize this lengthy example. Hayek passes from an aggregated model with one homogeneous representative individual in 1928 to a more disaggregated model with a multitude of individuals who are heterogeneous in the sense that each of them occupies a unique position in the system and may have different subjective preferences, in 1937. This procedure illustrates that for different problems we

need different models, each with their own simplifying assumptions and other methodological devices.[33] What links or unifies the models is the hypothesis of rationally choosing individuals, something they have in common with Walras' and Pareto's general equilibrium theories. In Hayek (1928) the problem was how to model the perceptions, the expectations and the equilibrium of the individual plan that is based on them, and the mechanism that may disturb this equilibrium, subject to the constraints Hayek had imposed on himself (such as the exclusion of the general price level as an explanatory factor[34]). For these purposes Hayek introduced the representative individual, stating that he intended to analyze the problem of the coordination between individuals—and hence the equilibrium of the economy—separately. He addressed this problem subsequently, in his 1937. Hayek's procedure is an instance of factualizing an idealizing model—that of the representative individual and what it means for him to be in equilibrium—through a generalization—from one representative individual to the coordination between all individuals or, what is the same, the equilibrium of the economy. That this concept of equilibrium can be generalized from economics to sociology is not an issue I want to pursue here.[35] Nor do I want to go into the role of rational choice theory in sociology; Cherkaoui has said illuminating things about this.[36]

EXPLANATORY SCHEMES AND THE METHOD OF DECREASING ABSTRACTION

Hedström, Swedberg & Udéhn (1998) raises the question why Popper's rationality principle is so little known among sociologists. I think an important part of the answer can be found in the piece of intellectual history I have just sketched: Popper himself failed to elaborate an internally coherent and complete methodology of social science. It was Koertge who did so. Her methodology employs the procedure that Hayek had used in his economics: the method of decreasing abstraction (MODA). In the final part of this chapter I will argue that sociology may benefit as much from this method as does economics. MODA, moreover, would be a natural complement to the methodology that Cherkaoui proposes and in which the rational individual plays an important part.[37]

A good point of departure for a decreasing-abstraction explanation is the representative individual.[38] The example of Hayek's business cycle theory shows how the representative individual and MODA are instrumental in explaining a particular collective phenomenon, the lack of intertemporal coordination. Hayek chooses a rational representative individual but alternative behavioural models are possible.[39] In sociology the device of the representative individual is rarely used—if ever. And why should it, indeed? Sociology is mainly interested in the relationships between individuals, which is exactly what the representative individual abstracts from. That is, of course, true—as far as it goes. I will argue that MODA makes it possible to solve the problem of combining empirical and theoretical strength. In addition, I will also try to show that MODA may serve to remedy causal incompleteness.

These tasks are made simpler by the fact that a detailed description of MODA and its application in sociology is available. It can be found in the work of Siegwart Lindenberg. Over the last 40 years Lindenberg has developed a research programme in which he gradually constructs the elements of a methodology that is capable of (1) explaining collective phenomena with the help of hypotheses about rational individual behaviour; (2) integrating sociology and economics; (3) integrating theory and empirical evidence in sociology; and (4) fruitfully employing rational choice theory in the explanation of aggregate phenomena. The parallels with Cherkaoui's work are obvious.

Lindenberg's exposition of MODA is based on the separation between analytical primacy on one hand and theoretical or explanatory primacy on the other.[40] By the former he means the domain of the intended explananda, which in the case of sociology and much of economics (particularly macro-economics) are aggregate or collective phenomena. In so far as the mechanisms that produce these phenomena are based on purposive human action, the explanatory or theoretical primacy lies on the individual level. It should not, however, be taken for granted that any theory about individual action is suitable for explaining aggregate phenomena. Most psychological theories, for instance, have been designed for an empirically adequate explanation of individual phenomena: their analytical primacy is the individual level. These theories also aim at generality in the sense of explaining many

apparently different individual phenomena. Usually, theories of this type are not general in a different meaning of the word: the capacity to explain aggregate phenomena.[41] In order to keep the two meanings distinct, Lindenberg introduces the terms individual$_1$ for theories that are intended to explain phenomena at the level of the individual, and individual$_2$ for theories about individual behaviour that are intended for explaining collective phenomena. Individual$_2$ theories are not primarily meant to provide accurate descriptions of individual behaviour; their main purpose is that of guiding and unifying a series of ever more accurate models of aggregate phenomena. In economics this role is played by rational choice theory. Such an explanatory "core" is made more realistic by supplementing it with a series of so-called bridge principles.

Let me make three observations. First, Lindenberg's distinction is a simplification of the more detailed scheme on page 125 of *Invisible Codes*. That scheme gives a typology of theories according to the type of problem (explaining individual and social patterns; this is what Lindenberg calls the analytical primacy) on the one hand and both the level of analysis (individual, individual-context, and structure) and the type of behavioural hypotheses (rational vs. others) that are used as explanantia (this corresponds to Lindenberg's category of explanatory or theoretical primacy). Second, it neatly captures the crux of Popper's and Hayek's antipsychologism. It makes it particularly clear that it does not imply a rejection of psychology. That would also be odd for scholars both of whom made important contributions to psychology and the philosophy of mind. What they do oppose is the use of individual$_1$ theories in the explanation of collective phenomena; individual$_1$ theories are legitimate and necessary for explanations of individual behavioural and psychological regularities. For the explanation of collective phenomena, however, we need individual$_2$ theories.[42] Using Lindenberg's terminology, Popper's and Hayek's antipychologism can be summarized by saying that it is a rejection of confusing individual$_1$ with individual$_2$ theories in the explanation of collective phenomena. Third, Lindenberg occasionally refers to Popper, particularly in his discussions of rational choice theory. As far as I know, he never mentions Hayek, even though that would have been natural for his presentation of MODA: objectively speaking,[43] Lindenberg's research programme is a direct descendant

not only of Popper's philosophy of science and his methodology of the rationality principle, but also of Koertge's elaboration and Hayek's use of MODA.

Lindenberg argues that MODA makes it possible to combine theoretical and empirical strength:

> [T]he method of decreasing abstraction [is] a means to integrate economic, sociological, and psychological lines of research without losing the analytical power of the economic approach or the descriptive advantages of the sociological and psychological approaches. As it turns out, this method also allows a stronger integration of model building with empirical research in general. (Lindenberg, 1992: 4)

So, MODA helps to solve the first problem on Cherkaoui's agenda, the integration of theoretical and empirical research in sociology. More in particular, it is a flexible tool that allows us to produce explanations that are "as simple as possible and as complex as necessary." (Lindenberg, 1992: 18). The method is a major instrument for linking simplified models that have great analytical power with more complex models that are more realistic description of empirical reality (ibid.). The description of the main features of MODA that Lindenberg gives (Lindenberg, 1992: 18–19), and which I will not repeat here, clearly shows that it is an elaboration of the central ideas that underlie Popper's and Koertge's methodologies and of the method Hayek uses in his economics.

TOWARDS CAUSALLY MORE COMPLETE EXPLANATIONS

Apart from being a natural complement to Cherkaoui's adoption of a rational representative individual, MODA has the added advantage of being instrumental in remedying causal incompleteness, which I have identified as Cherkaoui's second problem. It is commonly assumed that causes are described by the initial conditions of a deductive-nomological explanation. This is illustrated in the following simple example. Let's assume that "all men are mortal" is a law-like statement. The most

general explanation of John F. Kennedy's death follows the scheme: 1. All men are mortal. 2. J.F. Kennedy is a human being. 3. Therefore J.F. Kennedy is mortal. To this we only have to add the statement that all mortals eventually die. So, the, or rather, a cause of Kennedy's death is the fact that he was a human being. Of course, this is not a very interesting cause. More interesting is exactly how he met his death on November 22nd, 1963 in Dallas, Texas. In fact, there is no lack of theories about the specific causes of his decease. Each of them claims to supply some or all of the specific initial conditions that explain his death, or, what is the same, the causes that led to it. They give causally more complete explanations. How the various specific causes are related to the more general explanation of Kennedy's death ("he has died because he was a human being") involves invoking auxiliary hypotheses. This is exactly the procedure that is modelled in Koertge's scheme and followed in MODA explanations.

The same is true of explanations of collective phenomena: individual and structural causes can be integrated into a causally more complete framework that uses one principal type$_2$ theory, which by MODA is enriched by a series of initial conditions that may hail from type$_1$ theories. I use type$_1$ and type$_2$ to maintain the distinction between analytical and explanatory primacy while at the same time leaving open the possibility that they need not lie at the collective and individual levels respectively; the role of these levels may, for instance, be reversed. That is what the scheme on page 125 of *Invisible Codes* shows. It gives a classification of a variety of types of actually existing theories in the social sciences. Let us limit ourselves to the groups of theories that have the same analytical primacy but different theoretical primacies. None of these theories have exactly the same explanantia but their explananda lie at the same analytical level.

One might use this as a point of departure for setting up an explanatory experiment. Hypothetically, it is possible to start the explanation of the same collective or individual phenomenon (a specific invisible-hand process or some regularity at the individual level) using not only a theory about individual behaviour as the principal explanatory theory; one might also use a theory about the relationship between aggregate entities as the principal explanatory theory and gradually enrich or

factualize it by adding individual factors. This implies a generalization of MODA as Lindenberg presents it, one that incorporates several of the entries in Cherkaoui's scheme. So this generalized explanatory scheme is either based on a highly simplified model of individual behaviour, involving the representative individual (possibly, but not necessarily, rational) or on a simplified model in which collective phenomena are the primary causal factors. The important thing is that both models allow of a series of factualizations according to MODA.

The experiment would then consist in organizing a tournament between a collective and an individualistic explanation in order to discover which is the more satisfactory for a particular invisible-hand process or a particular individual phenomenon.[44] Of course, this is a very speculative proposal. In order to put it into practice, non-trivial theoretical and practical problems would have to be overcome. A first problem would be how to construct the explanatory (or type$_2$) theories and the series of ever more factual models in such a way that they can be *made* comparable.[45] That may be a daunting task. And even in case this obstacle and others could be overcome, it would certainly be too naïve to expect that the two competing processes of factualization lead to a shared model somewhere at an intermediate level of aggregation. As the example of Hayek's theory of the business cycle illustrates, every explanatory strategy has specific objectives, which guide the choice of specific idealizations, simplifications and factualizations. On the other hand, keeping a possible contest between theories with different explanatory primacies in mind[46] may help make these choices clearer, better argued, and even more rational.

NOTES

1. I thank Gianluca Manzo for his numerous valuable suggestions and references to the literature. What I have done with them remains, of course, entirely my own responsibility.
2. Cf. for example chapter 4.
3. For which Siegwart Lindenberg has coined the expression "Aggregatspsychologie." He calls it the enemy of the theoretician of collective phenomena. Cf. Lindenberg (1977: 60).
4. The work of Alan Kirman is an early and important exception. It takes into consideration both the rules of individual behaviour and the structure of interaction

and communication. This has given rise to a research tradition in economics that takes networks seriously.

5. Cf. Burt (1992: 2): "The unit of analysis in which structural holes have their causal effect is the network of relations that intersect in a player. The intersection is known by various names, depending on the context; it may be termed a role, a market, or a position in social structure. The players in which relations intersect are physical and legal entities: a person, an organization, or a broader aggregation of physical and legal entities. Where an intersection occurs, is merely an empirical curiosity; *causation resides in the intersection of relations*." (italics mine). Cf. Also Lindernberg (1995: 81).
6. Cf. Cherkaoui (2005: 117): "The error of inferring the social from the individual corresponds to that of inferring the individual from the social (the 'contextual fallacy') …".
7. More than by Weber. Hedström *et al.* (1998) raises the question of the authors who may have influenced Popper's methodology of social science. They mention Hayek only in passing.
8. Apart from the phenomenon of scarcity, two of the most well-known instances of these in economics are the Keynesian multiplier and the money multiplier.
9. For this, cf. Cherkaoui (2005), chapter 2.
10. It is part of his battle against essentialism, a topic I will not go into here.
11. For other influences than those discussed here, cf. Birner (2009 and 2014).
12. "In a way, "Economics and Knowledge" was an attempt to persuade Mises of the limits of the a priori approach!" (correspondence with the author, 17 July 1981). Cf. Also Kresge & Wenar (1994: 72), where Hayek speaks of "my 1937 article on the economics of knowledge, which was an attempt to persuade Mises himself that when he asserted that the market theory was a priori, he was wrong; that what was a priori was only the logic of individual action, but the moment that you passed from this to the interaction of many people, you entered into the empirical field."
13. Which he was already interested in. He had read work by Eugen von Philippovich (as he told Karl Milford—personal communication). Later, when he lived in New Zealand, Popper asked Harold Larsen and later Colin Simkin for help in explaining economics to him. Cf. Simkin (1993: 184).
14. Malachi Hacohen correctly observes that this shows the influence of Max Weber. Cf. Hacohen (2000: 489). He does not discuss Hayek's influence on Popper's methodology of social science.
15. In the only version that was available for quite a while, published as an extract in French as Popper (1967).
16. Cf. Koertge (1975 and 1979).
17. Cf. Hayek (1937: 48).
18. In all fairness to Hayek I have to add that his attempt to separate the two parts of economic theory into a non-falsifiable set of statements and another set that describes its empirical content is very similar to Lakatos's later distinction between the hard core of a research programme (the pure logic of choice in

economics) and a protective belt (the empirical statements about the way individuals acquire knowledge), which was proposed more than thirty years later. It is no coincidence that the methodologies both authors propose deal with disciplines that have a strong a-priori or tautological "feel" to them, economics (Hayek) and mathematics (the origin of Lakatos' methodology of scientific research programmes). This deserves more discussion than there is room for here.

19. Cf. also: "The world of Science might … be described as no more than a set of rules which enables us to trace the connections between different complexes of sense perception." (Hayek, 1955: 20). Hayek's philosophy of social science derives directly from the ideas he developed in his philosophy of mind. Cf. Birner (2012 and forthcoming).
20. "The origin of this idea may explain why I rarely stressed the fact that I did not look at the logic of the situation as a deterministic theory: I had in mind the logic of situational choices." (Popper, 1963: 181, quoted above).
21. Latsis (1972).
22. The first was the proposal of the "zero method" in *Poverty.*
23. Cf. Birner (2001).
24. Cf. Koertge (1975 and 1979). For the use of the term "factualization", cf. Birner 2001. Koertge does not use this terminology. She had written her PhD thesis on the correspondence principle under the supervision of Heinz Post (Koertge, 1969). She analyses idealizations explicitly in Koertge (1977).
25. See Koertge (1979: 91–2).
26. Koertge (1979: 89).
27. The domain of Koertge's auxiliary hypotheses is partly that in which later experimental economics has discovered and analyzed "anomalies". A corollary of Koertge' solution for making Popper's rationality principle coherent is that it also solves the problem of determinism that Latsis had signalled.
28. Cf. for this also Hedström *et al.* (1998: 356–358).
29. For the distinction between Hayek's analytical and associative work, cf. Birner (forthcoming).
30. Let me remind the reader that a central objective of Hayek's theoretical exercise is to show that relative prices influence economic processes and not the general price level.
31. Readers who are more interested in methodological than in economic solutions might skip the rest of this paragraph. The latter may want to consult the more extensive discussion in Birner (2000). I will be happy to send the original version in English to whoever asks for it. I hope the expert reader will forgive me the many simplifications I make for the purpose of this chapter. A detailed discussion can also be found in Klausinger (2012a and 2012b).
32. This is argued in Birner (1999). Cf. also Birner and Ege (1999).
33. Uncharacteristically, Hayek does not accompany this part of his economics with a detailed description of the method used.
34. Cf. also note 40.
35. Cf. Birner and Ege (1999).

36. For instance in chapters 3 and 5 of *Invisible Codes*. Cf. also Manzo (2013), who extends the analysis of the use of the rational-individual hypothesis into the domain of experimental economics.
37. And hence is in the spirit of Cherkaoui's thought. Cf. "what is true of sociology is also true of economics: it is not so much the unit of observation that matters as the way in which problems are simplified." Cherkaoui (2005: 122).
38. Cf. also Kirman (1992). I have already mentioned Kirman in the introduction as a pioneer in economics. In this seminal article he denounces the neglect by neoclassical economics of the structure of interaction between individuals. In subsequent work he went on to remedy it.
39. The use of the satisficing individual, which was introduced by Herbert Simon, is an example of an alterative hypothesis.
40. Lindenberg (1992: 7).
41. John Maynard Keynes uses "general" in this sense in the Preface to the German and French editions of his *General Theory*: the theory is general because it is a theory of the economic system as a whole.
42. I add: in conjunction with MODA, as the following discussion hopes to make clear.
43. Or, as Popper would say, at the level of World-3.
44. The contest may also be of heuristic value.
45. For observations about making theories comparable, cf. Birner (2002). Lindenberg (1996) also contains much that is relevant to this problem, and to the problems of integrating theoretical and empirical analysis and the causal incompleteness of many explanations in social science.
46. Hayek did so. One of his main purposes was to give an explanation of the business cycle that did not rely on the aggregate price level. An example of the contest I hint at in the text could be a very systematic comparison between the explanations of the business cycle in terms of aggregate, macroeconomic factors (the bulk of the theories both in the 1930s and nowadays) with an individualistic, microeconomic one such as Hayek's. Let me add that debates on the micro-foundations of macroeconomics have pursued different objectives.

REFERENCES

Birner, J. (1999) "Making Markets", in Dow & Earl, eds (1999), pp. 36–56.
——— (2000) "Nascita e struttura del programma di ricerca economico di F.A. Hayek", in Clerici & Rizzello (2000), pp. 268–299.
——— (2001) *The Cambridge Controversies in Capital Theory; A Study in the Logic of Theory Development*. London: Routledge.
——— (2002) "A Conservative Approach to Progress In Economics", in Boehm *et al.* (2002), pp. 65–88.
——— (2009) "From Group Selection to Ecological Niches. Popper's Rethinking of Evolution in the Light of Hayek's Theory of Culture", in Parusnikova and Cohen (2009), pp. 185–203..

——— (2014) "Popper and Hayek on reason and tradition", *Philosophy of the Social Sciences*, 44(3): 263–281.
——— (forthcoming) "Algunos problemas en los programas de investigación de Hayek y Popper: (des)unidos por el racionalismo", in Fernández forthcoming.
Birner, J. and Ege, R. (1999) "Two Views on Social Stability: An Unsettled Question", *American Journal of Economics and Sociology*, 58(4): 749–780.
Boehm, S., Gehrke, C., Kurz, H. and Sturm, S., eds (2002) *Is There Progress in Economics?* Cheltenham: Edward Elgar.
Burt, R.S. (1992) *Structural Holes. The Social Structure of Competition*. Cambridge (MA): Harvard University Press.
Cherkaoui, M. (2005) *Invisible Codes. Essays on Generative Mechanisms*. Oxford: The Bardwell Press.
Claassen, E.M., ed. (1967) *Les fondements philosophiques des théories économiques. Textes de Jacques Rueff et essais rédigés en son honneur 23 août 1966*. Paris: Payot.
Clerici, G. & Rizzello, S., eds (2000) *Organizzazione, informazione e conoscenza. Saggi su Hayek, volume I*. Turin: UTET.
Coleman, J.S. & Fararo, T.J., eds (1992) *Rational Choice Theory. Advocacy and Critique*. New York: Sage.
Dow, S.C. & Earl, P.E., eds (1999) *Economic Organisation and Economic Knowledge: Essays in Honour of Brian Loasby*. Edward Elgar.
Eichner, K. & Habermehl, W. eds (1977) *Probleme der Erklärung sozialen Verhaltens*. Meisenheim am Glan: Anton Hain.
Fernández, F. (ed.) forthcoming, *Karl Popper hoy*. Rosario: Editorial Juglaria.
Hacohen, M. (2000) *Karl Popper—The Formative years, 1902–1945: Politics and Philosophy in Interwar Vienna*. Cambridge: Cambridge University Press.
Hayek, F.A. (1928) "Das intertemporale Gleichgewichtssystem der Preise und die Bewegungen des 'Geldwertes'", *Weltwirtschaftliches Archiv* 28.
——— (1929) *Geldtheorie und Konjunkturtheorie*. Vienna: Hölder-Pichler-Tempsky.
——— (1937) "Economics and Knowledge", in Hayek (1949).
——— (1943) "The Facts of the Social Sciences", in Hayek (1949).
——— (1949) *Individualism and Economic Order*. London: Routledge and Kegan Paul.
——— (1952) *The Sensory Order. An Inquiry into the foundations of Theoretical Psychology*. Chicago (IL): Chicago University Press.
——— (1955a) "Degrees of Explanation", in Hayek (1967).
——— (1955b) *The Counterrevolution of Science*. The Free Press.
——— (1964) "The Theory of Complex Phenomena", in Hayek (1967).
——— (1967) *Studies in Philosophy, Politics and Economics*. University of Chicago Press.
Hedström, P., Swedberg, R. and Udéhn, L. (1998) "Popper's Situational Analysis and Contemporary Sociology", *Philosophy of the Social Sciences*.
Keynes, J.M. (1936) *Allgemeine Theorie der Beschäftigung, des Zinses und des Geldes*. Berlin: Duncker & Humblot.
——— (1939) *Théorie generale de l'emploi, de l'intérêt et de la monnaie*. Paris: Payot.
Kirman, A.P. (1992) "Whom or What does the Representative Individual Represent?", *Journal of Economic Perspectives*, 6(2): 117–136.

Klausinger, H.J. ed. (2012a) *The Collected Works of F.A. Hayek, Vol. 7, Business Cycles, Part I*. Chicago (IL): Chicago University Press.
——— ed. (2012b) *The Collected Works of F.A. Hayek, Vol. 7, Business Cycles, Part II*. Chicago (IL): Chicago University Press.
Koertge, N. (1969) *A Study of Relations between Scientific Theories: a Test of the General Correspondence Principle*. PhD thesis, University of London.
——— (1975) "Popper's Metaphysical Research Program for the Human Sciences", *Inquiry*, 18(4): 437–462.
——— (1977) "Galileo and the Problem of Accidents", *Journal of the History of Ideas*.
——— (1979) "The Methodological Status of Popper's Rationality Principle", *Theory & Decision*, 10: 83–95.
Kresge, S. and Wenar, L., eds (1994) *Hayek on Hayek: An Autobiographical Dialogue*. London: Routledge.
Latsis, S. (1972) "Situational Determinism in Economics", *British Journal for the Philosophy of Science*, 23(3): 207–245.
Lindenberg, S. (1977) "Individuelle Effekte, kollektive Phänomene und das Problem der Transformation", in Eichner & Habermehl (1977).
——— (1992) "The Method of Decreasing Abstraction", in Coleman and Fararo (1992).
——— (1995) "Complex Constraint Modeling (CCM): A Bridge Between Rational Choice and Structuralism," *Journal of Institutional and Theoretical Economics*, 151(1): 80–88.
——— (1996) "Choice-Centred Versus Subject-Centred Theories in the Social Sciences: The Influence of Simplification on Explananda", *European Sociological Review*, 12: 147–157.
Manzo, G. (2013) "Is Rational Choice Theory *Still* a Rational Choice of Theory? A Reply to Opp", *Theory and Methods/Théorie et méthodes*, 52(3): 361–382.
Miller, D. ed. (1983) *A Pocket Popper*, New York: Fontana Paperback.
Parusnikova, S. and Cohen, R.S., eds (2009) *Rethinking Popper*, Boston Studies in the Philosophy of Science, vol. 272. Berlin: Springer.
Popper, K.R. (1934) *Logik der Forschung*. Berlin: Julius Springer.
——— (1957) *The Poverty of Historicism*. London: Routledge & Kegan Paul.
——— (1957a) "The Aim of Science", in Popper (1972).
——— (1963) "Models, Instruments and Truth. The status of the rationality principle in the social sciences", in Popper (1994).
——— (1967) "La rationalité et le statut du principe de rationalité", in Claassen (1967), also retranslated into English and published as "The Rationality Principle", in Miller (1983).
——— (1972) *Objective Knowledge. An Evolutionary Approach*. Oxford: Clarendon Press.
——— (1994) *The Myth of the Framework: In Defence of Science and Rationality*, ed. by M.A. Notturno. London: Routledge.

6

METHODOLOGICAL INDIVIDUALISM AND THE NOTION OF LEVELS OF SOCIAL LIFE

Pierre Demeulenaere
University of Paris-Sorbonne

INTRODUCTION

Methodological individualism (MI) is a theory about human agency in a social world and about the articulation of the various entities constituting such a world. It is consequently a theory of explanation of social phenomena. It should not be seen primarily as a strategy to explain phenomena: its requirements about explanation are only the consequence of the way social reality is inescapably organised. In this paper, I will therefore not start with MI as a principle of explanation in the social sciences, but as a mapping of social reality. It includes the notion of levels: it is very common now to approach social reality through three notions (or metaphors) of levels, the macro one, the micro one, and the infra-individual level. Sometimes is added the notion of a meso level. These notions of levels are at the heart of the justification as well as the critique of MI. They have been one of Mohamed Cherkaoui's major interests (Cherkaoui, 2005). The characterisation of what is exactly meant by such a notion of level is therefore a major task for any theorisation of MI.

There are indeed today basically three ways of challenging the validity of MI. The three of them rest on an assessment of the "level" issue.

The first one is to assert that there exist laws at a macro level, which are not necessarily explainable on the basis of individual actions; therefore MI would be incomplete and false.

The second is to consider that the individual level of MI is too individualistic, and should be replaced by a more open theory of action which would have basically two dimensions: the actions should not be seen in an atomistic way, and they should involve the importance of the role of social settings in determining social behaviour (that is the importance of the macro level in determining the micro level).

A third attack comes from new trends in naturalism that stress infra-individual determination of behaviour, and consider that a unique reference to the individual level disguises the force of unconscious motives of actions and of neuronal determinations.

I will argue in this paper that those three critiques are inaccurate, incomplete, and do not really challenge the core idea of MI (as expressed by Max Weber) and its consequences in terms of explanation. My contention is not that those critiques are completely irrelevant; it is that they are based on incomplete, inaccurate of false assumptions about MI. One of the reasons of such a misunderstanding is the reification of so-called levels, and a misplacement of their role.

1. MI AS A THEORY OF THE UNIQUE ROLE OF ACTION IN SOCIAL LIFE.

The main idea that can be said to characterise MI is that, in social life, social entities do not act by themselves, only individuals (i.e. human beings) do so. Since only individuals can act, methodological individualists consider that social entities (like states, institutions, organisations) do not act by themselves, but need individuals to be set in motion. This rather simple and trivial (and in my view inescapable) idea is however compatible with two other major propositions.

First, institutions and social entities can, under certain conditions, be assimilated to actors. However, in the tradition of MI, it remains that

such corporate actors need individuals to be set in motion; they just cannot act without the intervention of individual actors.

Second, social entities can have an "influence" on individual behaviours, although this notion of influence can be interpreted in various ways, more or less causal. It depends on what is to be considered as a causal factor and on the nature of the social entities that are taken in consideration when we speak of an influence or of a causal role.

I will try to present the logical consequences of the main idea developed by MI, that only individuals act, in order to show that it is not challenged by MI critiques. I easily acknowledge that some eminent authors in this tradition have defended narrow versions of MI, which add supplementary (and unnecessary) hypotheses to the basic proposition of MI and go beyond the core idea. When such narrow versions are at stake, some of their critiques are relevant. But, as I have said, I will focus on the core idea and its logical consequences.

If we set aside the difficult problem of non-human animal behaviour, that should however be addressed by sociological theory, it can be said that in social life only individuals act. This is clearly stated by Max Weber:

> Action in the sense of subjectively understandable orientation of behavior exists only as the behavior of one or more *individual* human beings. (Weber, 1978: 13)

> For still other cognitive purposes—for instance, juristic ones—or for practical ends, it may on the other hand be convenient or even indispensable to treat social collectivities, such as states, associations, business corporations, foundations, as if they were individual persons. Thus they may be treated as the subjects of rights and duties or as the performers of legally significant actions. But for the subjective interpretation of action in sociological work these collectivities must be treated as solely the resultants and modes of organisation of the particular acts of individual persons, since these alone can be treated as agents in a course of subjectively understandable action. (Weber, 1978: 13)

This fundamental proposition of MI purports that there is an active social life only because there are individuals who act. Weber speaks here of intentional behaviour. But it could be affirmed, more simply, that human behaviour, intentional or not, is the active principle that sets in motion social life. This is a descriptive proposition about what is responsible for motion in social life. It is not directly an ontological proposition about what exists (or does not exist) in the social life, nor is it a position about explanatory possibilities in the social sciences, although this fundamental proposition has important consequences regarding those two matters.

Once said that only individuals act, three sets of questions derive from this elementary proposition: the three of them refer to the description of so-called "levels" in social reality and of their mutual relationships.

The first set is about the characterisation of the so-called individual level, that is action as such. It can be misleading, since the notion of "individuals" can refer to human beings as actors, or more specifically to a possible "individualism" of such actors. What does indeed an individual action exactly mean? Does it imply that individuals are separated from each other and do not interact? Does this mean that they are supposed to act in a certain way (a selfish or individualistic one for instance)? And, in particular, are individuals supposed to be rational? If it is said that they are rational, does this imply that they are not influenced by institutions and by social settings (and therefore the so-called social level) in the course of their action?

The second set is about the infra-individual level, and its role in the determination of social behaviour. Does MI exclude an infra-individual description of action, and should the introduction of such a level lead to a MI dismiss?

Finally, the third set of questions is about the role and characterisation of the so-called macro level, with three major problems. Does the assertion that only individuals act imply that only individuals exist and that macro-level entities do not exist as such? Does the assertion that only individuals act imply that any explanation of the macro level should be only based on the micro level? And, finally, does the reference to the individual level in order to explain the macro level imply

that there are no laws at the social level (this question will not be treated here)?

I will try to discuss rapidly these dimensions on the only basis of the core idea of MI, expressed by Max Weber, that only individuals act.

2. THE CHARACTERISATION OF THE INDIVIDUAL LEVEL

In this section I will argue that MI does not imply individualism, nor a reference to individuals separated from social context or institutional situation. Consequently, the description of individual action is always linked to a social context, and the micro level involves a reference to the macro level.

Stressing the uniqueness of individuals as the sole possible actors in the social world does not imply any kind of "individualistic" position, as Weber clearly insisted on. MI is frequently associated with neo-classical economics in which social outcomes (like market price equilibrium) are derived from separate individual decisions. In this scenario, individuals are separated from each other since, in the model, the formation of price equilibrium does not require their formal agreement to reach this point. Moreover, in the economic tradition, individuals not only do not cooperate and do not discuss together in order to produce equilibrium prices, but are selfish: they try to achieve a maximisation of their personal utility. These hypotheses are at the core of economic neo-classical construction. They can be found equally in the theory of games, where actors similarly do not cooperate, in the sense that they do not discuss together the situation in which they have to play and the way to deal with it, and do not have initially a cooperative attitude. On the basis of such behaviours, specific social outcomes can be seen to emerge under certain conditions: for instance price equilibrium or social norms.

However those two dimensions (absence of direct interaction and selfishness) are not intrinsically necessary to MI as such, either in theory or in practice. The idea that only individuals act has no logical implication on the nature of their relationships: they can indeed interact and deliberately cooperate and set up social norms. They can also have altruistic behaviours. Sociologists have frequently negatively considered MI

as an "atomism" involving separate actors, who are moreover selfish. But this is not necessarily the case. It is true that Menger (1996) for instance referred to atomism, insisting that money, when it emerges as a social institution, does not need to be the result of an agreement between actors. It derives as an unintended consequence from separate 'atomistic' decisions.

Generally speaking, it can be useful, sufficient, or relevant, in specific social situations, to suppose that actors act separately and do not cooperate: it is the case for instance in the stock exchange or, more generally, in the ideally conceived markets of neo-classical economics. It can be shown, moreover, that in certain theoretical types of situation (in particular prisoner's dilemma situations) separate selfish individuals can be led to adopt cooperative norms in order to satisfy their interests. But this does not mean that all norms derive from such situations and such behaviours. That type of description corresponds to the fact that, in many social situations, individuals in effect do act separately, in a selfish way moreover. On the basis of such descriptions, we can have the elements from which can be derived social outcomes. It should be added that such scenarios presuppose themselves the antecedent existence of social settings: the market, or the specified conditions of the interaction in the theory of games.

But the idea that only individuals act does certainly not imply that it is always the case that individuals do not interact or act in a selfish manner. Interaction can be found even in economics, based on selfish motives, when for instance actors set up monopolies or oligopolies. This is clearly the result of an interaction among individuals. Therefore, the specific economic tradition that displays separate actors is not to be conflated first with the whole of economic analysis and second with MI.

As it has already been suggested, the idea that only individuals act does not imply that they act in a selfish manner. When J.S. Mill (1988), against the tradition of Comte, insisted on the importance for social phenomena to be based on the action of individuals, and stressed the importance, for economics, to suppose a selfish behaviour, defining the *Homo economicus* figure (although he did not use the term), he certainly did not conclude that all men in all circumstances acted in a selfish manner aimed at increasing personal wealth. On the contrary

he clearly stressed the possibility of moral and altruistic behaviours. However, methodological individualism is frequently reduced to an economic analysis of selfish behaviour. Once again this cannot be derived from the core idea that only individuals act.

Moreover, the very idea of selfishness and of satisfaction of interests is relatively ambiguous in the economic tradition of action. On the one hand, starting with Mill, the figure of a narrowly defined economic man pursuing the increase of his wealth is designed in order to analyse economic markets. But on the other hand, very soon, the complexity of interests, the variety of preferences, and the fact that the notion of utility is a fundamentally vague one became preponderant in economic analysis: utility depends indeed on various psychological patterns on the one hand, and on social norms on the other hand. Therefore, Pareto first introduced the notion of ophelimity (utility based on individuals choices) and afterwards utility in economics became a formal notion derived from preferences. Preferences are diverse and can be linked to social norms. For instance smoking or taking drugs can be seen favourably or unfavourably on the basis either of personal tastes or of social prescriptions. There is therefore in such cases no intrinsic notion of "cost" or "benefit" that is at the bottom of the decision. The decision is itself what allows us to speak of costs or benefits, and they can be diverse for different persons or for various social settings: for one it is costly to take drugs (to drink alcohol for instance), for the other it is costly not to take them. Therefore, in the domain of economics, the theory of action linked to MI is fundamentally ambiguous: it rests on the idea that actors tend to maximise their benefits or minimise their costs, but since costs and benefits do not represent an unambiguous universal signification, the theory of action is not really predictive or effective of definite actions and choices (Sen, 1977; Demeulenaere, 1996).

Therefore, the distance between an economic model of social behaviour, and a sociological model of behaviour is strongly reduced (Akerlof and Kranton, 2010). Sociologists have generally attacked MI on two grounds: first, the idea that individuals are not selfish persons acting on the basis of a pure calculus of selfish interests; second that individuals do not act in a social vacuum but depend on social norms and institutions that constrain them. This second aspect will be further

developed below. But it should be straightforwardly noted that those two critiques do not attain the core of the idea of MI; moreover, it does not attain the core of the economic version itself of MI, since interests, preferences, and choices can rest on social norms: therefore social aspects can be found in the very heart of economic decisions and the question is thereafter whether there is or not a theory of decision devoid of any social aspect (and in particular social norms).

This leads to another related question: should the theory of individual decisions be taken from a naturalistic stance, where the elements of the decision precede social dimensions, or should it be considered as inherently social, in the sense that there would be no natural, stable and transcultural regularities in the determinants of the choices beyond the observed social and cultural variety in behaviours? This question can be reformulated in a way that involves the notion of level: does individual decision always precede the formation of the social level? This question is related to the scope and significance of rationality, in individual decisions, which can be seen either as a natural feature of the human species, or as a social norm prevailing in particular historical settings.

The idea that only individuals act has by itself no logical consequence regarding the question of cultural and social variation. Three positions are compatible with this core idea: one that would stress that all behaviours depend on a natural (namely psychological) basis, rational or not; another that would insist on the overwhelmingly cultural dimension of all individual behaviours, or the social position of the actors; and finally one that would consider that the two dimensions are always interconnected: the natural features of the human species lead to a diversity of social and cultural situations, since the main distinctive natural feature of the human species is, precisely, its aptitude to culture; this cultural diversity, though real and decisive for individual decisions, is not however disconnected from general natural traits of human behaviour (in particular emotions), and should tentatively be interpreted in their terms. It is likely that this third option is the right one (Haidt, 2012), and it is linked either to a theory of rationality or to unconscious motives (or to both).

Thus the debate about the rationality of action intervenes in the more general frame of the characterisation of action and behaviour. MI

is largely linked to (and criticised for) the idea that individuals act in a rational way. But here again, there is no necessary link between the idea of individual action and that of rational action. It is perfectly understandable that individuals can act irrationally (Elster, 2007). However, this depends on the notion of rationality we refer to and on its status.

There are five basic notions of rationality in the social sciences that can be classically linked to MI. All of them can be related to Popper's "problem-solving" notion.

The first one is to consider that individuals are rational whenever they act in an intentional way. The intentional decision to act corresponds to the reason someone has to act. An irrational behaviour would be then to act on the basis of non-intentional causes. For instance, Kahneman (2011) reports a study about parole judges observed in their decisions. In this example, individual judges tend to be more severe when they are hungrier: clearly in this case non-intentional and unconscious causes affect intentional decisions. This means that intentional motives do not have the sole influence in individual decisions; however it is still individual decisions that are at stake, partly determined by unconscious trends. Intentional action can moreover be linked to and dependent on social norms and in that case it includes social level dimensions.

The second notion of rationality is linked to the notion of consistency or transitivity of choices. The rationality is not related to the contents of individual choices as such, but to the fact that individuals are consistent in their preferences. This is the official position of neo-classical economics, where rationality does not involve any kind of "right" decision, but only the fact that people do not take inconsistent decisions. Irrationality here would be to have inconsistent choices. But again irrationality regarding consistency of choices would affect individual behaviours and herein would be compatible with MI. Here again, the transitivity of choices implies that the content of the choice is not part of the rationality dimension, and therefore depend on other dimensions (psychological or social). When they are social, they include a social level dimension.

The third notion of rationality links it to an instrumental dimension of action, where an actor chooses the adequate means to reach an end. The rationality derives from an ability to select the correct means,

which requires that a normative notion of correctness be here implied. It should be noted that although instrumental rationality is widely associated with standard economics, it can be separated from it from two opposed points of view: first, as was noted by Weber, instrumental action does not necessarily imply any kind of economic motives. Second, the transitivity of preferences, which is the core of neo-classical theorisation of economic action, does not in fact involve a notion of instrumental action. Here again, instrumental rationality presupposes and involves motives that can be related to a social level.

The fourth notion of rationality is clearly the most widespread in the social sciences, although it suffers major ambiguities (Demeulenaere, 1996) and can be the subject of deep criticism. It is the idea that people act rationally whenever they efficiently pursue their self-interest (Coleman, 1990). This is a much narrower conception of rationality, which does not insist only on the means to achieve an end, but on the ends of action: namely self-interest. In this scenario there is however a hesitation between two alternatives: either the preferences are supposed to be the same (like a preference for an increase of wealth), or they can be dissimilar (like smokers' or non-smokers' preferences). This theory is strongly normative since it reduces rationality to the choice of self-interest: this can be refused from a philosophical point of view. However, here again, if it appears that actors are not rational in this sense, that is if they do not behave out of strictly selfish motivations, it does not mean that it is not individual actions that are under consideration. It can be said in any case that self-interest is widely associated with social norms, for instance the way in India today many parents prefer to have boys instead of girls in order to avoid having to bear the price of the dowry.

The fifth notion of rationality is much rarer in social sciences, as a consequence of the influence of positivism: it extends its scope to the determination of right choices that can include not only correct beliefs but also normative deliberations. Boudon (2011) has advocated the reintroduction in sociology of this classical meaning of rationality in philosophy. This notion is always context dependant, and therefore linked to a social level.

In sum, whatever notion of rationality is selected, an irrational behaviour is by definition possible (otherwise it would be pointless to

determine the conditions of a rational action); but there is no reason to consider that irrational actions are not individual actions, or at least individual behaviours, since the notion of action can involve intentionality. Therefore rational behaviour is not consubstantial to MI, even if there are reasons to consider that a rational interpretation of behaviour should have priority over an irrational one, for one main reason, clearly expressed by Weber (although it is contested by Elster, 2010): when a rational behaviour is possible, it can be expected that an individual will adopt it since it is part of the normative idea of rationality that anyone should adopt the correct rational behaviour, when it is available. For instance, one should think that two and two make four, and not five. Therefore, if one adopts an irrational behaviour, we need a further investigation to understand why it is the case.

Naturally, from a more general point of view, we need a theory to explain why the human species is endowed with such a dimension of rationality, with its normative pressure, and what are its contents and properties. We need to explain rationality, as much as we need to explain irrationality at a meta-level, or from an evolutionary perspective. But, for social sciences explanations, on the basis of a given theory of rationality (and whatever are the origins of rationality), a reference to rational behaviour does not raise in principle supplementary questions about why someone adopts a rational attitude, whereas the evidence of an irrational behaviour raises the question of why the rational choice, when available (it is not always the case that a rational choice can be defined regarding a specific issue, and, when conceivable, is available to the information of the actor) was not selected. This is certainly not intentional fundamentalism (Ylikoski, 2012) but a consequence of the normative dimension of rationality.

I have so far indicated that the principal idea of MI, the unique role of actions in the movement of social life, has no logical consequence on the characterisation of individual actions. I want to consider now the consequence of this idea to the relation between individuals and their social environment, that is the individual level and the social level.

From its beginnings, MI has been attacked under the charge of "atomism", for instance by Durkheim, on the ground that in this scientific tradition not only individuals do not interact, but are completely

free of any social determination or influence. This critique cannot be accepted for two reasons. First, all majors MI sociologists, as well as philosophers like Popper (1966) or Watkins (1957), have accepted the importance of social situations as necessary to the interpretation of individual behaviours. It is the case for instance of Homans (1967). Thus when Wippler (1978) coined the term of "structural individualism", to stress the role of social situations in individual behaviour, it has been largely accepted, for instance, by Boudon and Bourricaud (1990) who explicitly quote this concept to describe MI. There is therefore no new thing regarding MI in introducing "structural individualism". The important issue for MI is not to deny such an influence, such a presence of the so-called "social level", but to understand the way it is exerted on actors, which occurs throughout various and heterogeneous ways. Therefore it is not correct to describe MI as a refusal of the role of social situations on individual behaviour as does Lars Udehn (2011). Second, the alternative is not between socialised actors unilaterally depending on social structures, and non-socialised individuals deprived of any kind of interpersonal relations and independent of any social influence. This is an untenable caricature. In the tradition of MI, people belong to social situations that exert influence on them: this is clearly the case for instance in Coleman's famous theorisation of social action (Coleman's boat), where individual actions are understood only on the basis of their social environment.

However, three things must be added in this respect:

— In the tradition of MI, social factors certainly exert an influence on individual actors, but individual actors are the ones who make the decisions, limited and constrained by these social influences. This means that social factors do not act by themselves, but individuals are constrained by and respond to social factors. It is therefore important to clarify and explicate the way by which social factors "have a role" in individual decisions.
— When they act within the limits of a social situation or react to a social situation, people can be sustained by clearly "natural" propensities, like a tendency to have sexual desires, or, alternatively, by clearly social dimensions (like the preference for a certain style of

food) they have previously adopted and sometimes "interiorised" (like prejudices about what is the right kind of sexual desires one should have). There is always a certain room for decisions, even in very constrained situations. It can thus be the case that there is some kind of "interiorisation" of norms or habits: this idea is accepted for instance by Coleman (1990). It implies specific mechanisms that bridge social settings and individual behaviour through psychological (and neuronal) adaptations, which are not necessarily clear and well understood today. For instance people raised in a specific language tend to have the accent of this language when they speak another one.

— However, a specific feature of MI is to try to interpret the emergence of social institutions and social situations on the basis of prior decisions and previous individual moves. For instance the design of a specific keyboard (of the QWERTY type) makes all people constrained by the way it is to be used, although another layout could be more efficient. But we also know the reasons for which such a design was selected for specific purposes.

As a general conclusion of this part, I would like to stress one major point: there is no consideration in MI of actions that are not referred, more or less explicitly, to the social situation. Therefore, the individual level is not considered independently of the social level, either because an individual decision always occurs within the frame of a social situation, or more strongly, because any individual decision involves prior social beliefs or norms that were adopted in the course of the individual life.

3. THE CHARACTERISATION OF INFRA-INDIVIDUAL LEVEL

The idea that only individuals act has been deployed by methodological individualists against holists, or structuralists, who evoke the action of social wholes. Today it would be reaffirmed against theorists who claim the existence of a social causation which does not involve individual actors (Kincaid, 1996). MI has been challenged in return, more recently,

by theorists who stress the importance of infra-individual determination of behaviour, either at a psychological level or at a neuro-physical level (Sperber, 2011). They underscore the existence of unconscious motives (or determinants of action), and interpret their role as the necessity to go to a deeper level of determination of action, which is therefore named infra-individual. Consequently, there are social phenomena, based on unconscious motives, for instance the diffusion of beliefs, which do not depend on individual deliberations.

There are two different things involved in this argumentation. First, as has been already said, there is no reason that MI should refuse the existence of unconscious determinants of action that are partly involved in individual behaviour (which is therefore partly intentional, partly non intentional). To acknowledge their existence does not lead to a refusal of the basic idea that only individuals act, even if this raises the complex problem of the relations between the conscious and intentional dimensions of behaviour on one side, and the unconscious and unintentional ones on the other. Therefore, whenever diffusion phenomena exist that do not depend on conscious choices, this does not mean that it is not through the means of active individuals (with their psychological characteristics) that those phenomena intervene. It would for instance be meaningless to refer to beliefs without referring to individuals who support these beliefs and whose psychological mechanisms are responsible for their diffusion.

Secondly, unconscious motives and subintentional determination of behaviours are frequently associated with an "inferior" level, which should be taken as causally responsible for actions, preferably to the individual level. But this is a misleading argument. It rests on the idea that the inferior level has no role in the determination of intentional behaviours themselves. But there is no basis for this. If one accepts (as promoters of a sub-intentional determination of action clearly do) that unconscious behaviours are determined by psychological or physical mechanisms that lie beyond the scope of individual awareness, it can be said, and it is in fact necessary to say, that conscious individual decisions and deliberations rest themselves as well on such sub-intentional mechanisms (that determine notably the possibility and strength of rationality) (Damasio, 1999). They are equally concerned by the mind body

problem. Therefore, a reference to psychological or unconscious motives does not straightforwardly lead to a sub-intentional level, since rationality itself necessarily rests on psychological and physical determinants. The infra individual level concerns rationality as well as irrationality: but to discover the biological mechanisms that are involved whenever a behaviour is engaged, rational or irrational, conscious or unconscious, is not the task of social sciences as such. It would be indeed very useful to have (which is not the case by now) a biological description of the neuronal mechanisms that are responsible for the development of rational choices or of irrational decisions, more or less linked to social norms. That would be a decisive contribution to the understanding of such decisions. But the social sciences have a specific role, which would not be replaced by such an illumination: to understand how social structures affect individual decisions and how individual decisions produce social outcomes.

Similarly, social norms that are displayed in a given situation have to be compatible with the general natural properties that characterise the human species. Therefore, any particular norm that can be found in a very particular culture (like drinking hot water in China instead of cold water), has to be compatible with the general features of human behaviour, which allow the formation of habits that incline either for drinking cold water or for drinking hot water. There are other norms that are universal, and can directly be interpreted in terms of natural selection, for instance the taboo against incest (Turner and Maryanski, 2005). But it is inaccurate to say that one type of norm (the universal one) has a biological basis, whereas the other one (the particular one), being cultural and variable, is not biological: a particular cultural norm, such as drinking hot water and being accustomed to that, necessarily depends on the general features of human biology that allow such behaviour (as well as the opposite one, a preference for cold water and a disgust for drinking hot water). The specific role of social sciences is to describe and interpret the emergence and variation of such norms and behaviours, universal or particular: they all depend on a biological basis, but are all linked to social situations.

4. THE CHARACTERISATION OF A MACRO LEVEL.

So far, following Weber, I have described MI as a theory asserting that, in social life, only individuals act. Consequently, it is not directly an ontological position about the existence of individuals or of social facts. The insistence on action is not an insistence on existence. However, it poses the problem of the relation between individuals and social entities, classically named and analysed as the micro/macro transition (Little, 1998; Cherkaoui, 2005). It involves the determination of what the social entities are and the nature of their link to individuals, that is, once again, the issue of so-called "levels" in the architecture of social life. Those problems are related to ontological questions about what is considered to exist in social life.

When we refer to the micro/macro articulation in social life, two different positions are possible and should be carefully distinguished.

(1) One is very general, and will consider that the micro elements are the parts of any set which include them. For instance, administrative entities are usually hierarchically dependent on other superior administrative entities (cities, regions, states, supra-national organisations). In this case, each level can be named and treated as a quasi-actor: it can be said for instance that Russia and the United States are players on the global scene. Therefore, micro and macro entities do not refer to specific levels, but to relative positional situations: for any macro level, there is generally a micro level. And when, between two levels, one being named micro, and the other macro, it is possible to localise an intermediate one, it can then be named a "meso" level (Turner, 2003). But this meso level could either be seen as a macro level or as a micro level from other relevant points of view. Therefore, there are no specific substantial meso levels as such, they always depend of a given positional relation between micro and macro levels.

However, two caveats should be added to this first relational position. First, since Russia and the United States, for instance, are stable long-term existent countries, they clearly have properties that are stable beyond individual decisions, which make them look like a quasi actor. For instance, they have a size that does not depend on actual individual decisions and which exerts constant constraints on the succession of

people who are in charge of these countries. However, even if it is not meaningless to speak of a Russian or an American strategy, on the long run, beyond actual individual decisions, due to the existence of those lasting properties, in fact, any interpretation of actual decisions by one of the states necessarily involves singular individuals (presidents, ministers, various officials and so on, constrained by the relatively permanent properties of the State they represent; moreover those properties do change, and the Soviet Union is not the Russian Federation). Quasi-actors are precisely not actors in the same way as individuals are. Quasi-actors need singular individuals to be active, so that there is only one privileged level that allows sound interpretations of the behaviour of quasi-actors. In this perspective, the fundamental relevant micro level, for social sciences, is the individual actor level.

Second, it would not be correct to describe a relation of "emergence" between the various levels: States are clearly not emergent regarding cities, although it can be said, from an administrative point of view, that cities belong to a State, and are parts of it. Moreover, the same elements can be part of various social sets, which do not overlap. For instance, an individual can be a member of a football club, and the club be part of a league in which countries as such play no role. The football player can be also a member of a club which belongs to a country which is not his own. There is no such thing in social life as a unique hierarchy of levels where superior levels emerge from inferior ones. Therefore, the analogy with the relation between the infra-individual level and the individual level is clearly misleading: it is not at all the same type of relation, since it cannot be said that the social emerges from the individual in the same sense (however problematic it is) as the individual mind would emerge from its body determinants. And this point is different from the multiple realisation of a macro level on the basis of plural alternative sets of micro elements.

(2) This is why, although it is possible to conceive the micro/macro relation as a purely positional relation, where no specific level is privileged, and where institutions can often be treated as quasi-actors, it is however important for MI to reaffirm that quasi-actors necessarily depend on singular actors to be active. And this leads to the second and more important meaning of the micro/macro relation: macro level

institutions, organisations or rules are necessarily referred to singular actors (embedded in social situations) to be understood. Individuals and organisations cannot be treated as quasi-actors, or as actors in the same interchangeable way.

What are then the social facts corresponding to the macro level? There are many words that are commonly used, the main of them being "collective", "structural", and "system" level. The micro/macro relation is moreover frequently interpreted as a whole/parts relation, following Durkheim. All this is relatively imprecise and refers to various positional or substantial realities, which pose different ontological problems. I will not describe here the way every particular author has mapped those social dimensions, nor will I develop an extended discussion about their ontology. I will only suggest (Demeulenaere, 2011) that there are basically three types of social facts regarding an individual level. The social level, or macro level, is therefore not a homogenous one and should not be treated in a general unspecified way. Consequently, there is no homogenous micro/macro link.

First, the macro level refers to the distributional properties of an aggregate, for instance a mean. We can select for example at random the people who enter a shop and determine their average age. In this case, it can be said that there is a "whole" since we select a definite number of persons, but there is certainly no such thing as an emergent whole regarding the parts. The constructed measure of the mean does not correspond to the properties of a whole as such, in the same sense as the age of one person is a property of this person (although this property is by no way a simple property, since it involves various institutional considerations about the appropriate way of measuring the age). The notion of aggregation can be here misleading. If we add up a series of stones, we have an aggregate, which has specific properties (for instance it might have a pyramidal shape). The aggregate has emergent properties (height, length, and so on) which the individual stones do not have. As opposed to this, when we measure the average size of the stones, we do not refer to this whole as such, but to the individual stones taken together. So the aggregate in this case can correspond to two different things.

A second meaning of the macro dimension corresponds to the relative positions of a series of elements: for instance an individual is taller

than another one, or richer, or more powerful etc. Here again, it could be possible to speak of a whole, or of a level, but these are metaphors that have no substantial basis: there is no emergence of a specific whole or of a level, beyond the set of elements that are envisaged from the point of view of their relative positions.

Finally, common rules that are followed (or rejected) by individuals represent a macro dimension regarding the actors who set up, accept, refuse or depend on them. These rules can include norms and institutions. Sometimes norms and institutions give birth to quasi-actors, who do not appear in the two previous dimensions of macro levels. In this case, it is easier to speak of a group level, since people may have a sense of belonging to the whole. This is why Searle's social ontology is focused mainly on this type of situation. However, the notion of a "superior" level is here ambiguous since rules, norms, institutions and organisations depend on actors for their life. They can be said to emerge as a social novelty, but they can disappear if nobody follows them. But even when people have a sense of belonging to groups corresponding to norms, it cannot be said that we are in a complete whole/ part relationship, because: (1) for any given group, the properties of the individuals who belong to it are not reducible to the properties of the group (that is the norms and institutions that constitute it); (2) some of the rules characterising a group can be detached from a specific set of individuals and referred to another one.

All those macro dimensions being deeply heterogeneous, they inevitably pose heterogeneous ontological problems about the notions of "levels", which will not be discussed here. The idea that only individuals act has no direct implication regarding the ontology of social facts, since it is clear and indisputable that there are existent things which do not act. The main position of MI is about the fact that only individuals act, not about the existence of social facts. The ontological problems depend on the notion of what an existent being is, and various positions will be possible regarding this. An organisation can be said to exist, even if it has not the physical properties of an individual. A measure exists, even if the mean to which it refers does not exist from a physical point of view. The debate is not so much whether social facts exist or not (it would be absurd to consider for instance that United States do

not exist), but what type of social facts exist, and what is their mode of existence.

CONCLUSION: EXPLANATION AND LEVELS

The main proposition of MI concerning the fact that only individuals act has consequences for the relevant explanatory strategy in social sciences. Globally, it is said that social facts should be explained by reference to individuals who set in motion social life. In particular, correlations between social variables (for instance gender and academic success) have to be explained by individual actions and decisions, in a given context (Boudon, 1974). But the actual requirements for relevant explanations depend on the nature of the considered social facts and of their relation to individuals. The relation that exists between a series of individuals and a collective measure is not the one that exists between a set of individuals and the norms they adopt or refuse. In the first case there is no social novelty to be explained, and the collective measure logically derives from the individual measures. On the opposite, in the second case, norms represent a social novelty regarding individuals, and are to be explained genetically by their beliefs and decisions, in a given situation (which already includes social norms).

The consequence of this is that although the requirement of MI is to insist on individual actions, since they are embedded in a social context that plays a role in their decisions, we cannot say that the macro level is fundamentally explained by a micro level in a reductionist move: since a specific social context is necessary to interpret the individual decisions, one dimension of the macro level is necessarily involved in the explanation of another dimension. However, a given context, locally admitted for the purposes of a definite explanation, can be, from another point of view, the focus of explanatory efforts.

Therefore two positions should be rejected: one that would say that MI requires to refer only to a micro level in order to explain macro level phenomena, since any explanation actually involves macro dimensions. And second, it would be incorrect to accuse MI of intentional fundamentalism, since any explanation involves a reference to social settings and necessarily includes unconscious dimensions of behaviour.

REFERENCES

Akerlof, George A. and Kranton, Rachel E. (2010) *Identity Economics. How our Identities Shape Our Work, Wages and Well-Being*. Princeton and Oxford: Princeton University Press.

Boudon, Raymond (1974) *Education, Opportunity and Social Inequality*. New York: Wiley.

——— (2011) "Ordinary rationality: the core of analytical sociology", in P. Demeulenaere (ed.), *Analytical Sociology and Social Mechanisms*. Cambridge: Cambridge University Press, pp. 33–49.

Boudon, R. and Bourricaud, F. (1990) *A Critical Dictionary of Sociology*. Chicago: University of Chicago Press.

Cherkaoui, Mohamed (2005) *Invisible Codes: Essays on Generative Mechanisms*, Oxford: The Bardwell Press.

Coleman, James S. (1990) *Foundations of Social Theory*. Cambridge and London: The Belknap Press of Harvard University Press.

Damasio, Antonio (1999) *The Feeling of What Happens: Body and Emotion in the Making of Consciousness*. New York: Harcourt Brace.

Demeulenaere, Pierre (1996) *Homo oeconomicus, Enquête sur la constitution d'un paradigme*. Paris: P.U.F. "Quadrige" new edition, 2003.

——— (2011) "Introduction", in Pierre Demeulenaere (ed.), *Analytical Sociology and Social Mechanisms*. Cambridge: Cambridge University Press, pp. 1–30.

Elster, Jon (2007) *Explaining Social Behaviour. More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

——— (2010) *L'irrationalité. Traité critique de l'homme économique*, II. Paris: Éditions du Seuil.

Haidt, Jonathan (2012) *The Righteous Mind. Why Good People Are Divided by Politics and Religion*. New York: Vintage Books.

Homans, George C. (1967) *The Nature of Social Science*. New York: Harcourt, Brace & World.

Kahneman, Daniel (2011) *Thinking Fast and Slow*. New York: Farar, Straus and Giroux.

Kinkaid, Harold (1996) *Philosophical Foundations of the Social Sciences: Analysing Controversies in Social Research*. Cambridge: Cambridge University Press.

Little, Daniel (1998) *Microfoundations, Method and Causation: on the Philosophy of the Social Sciences*, New Brunswick, NJ: Transaction Publishers

Menger, Carl [1883] (1996) *Investigations into the Method of the Social Sciences*. Grove City: Libertarian Press

Mill, John Stuart [1843] (1988) *The Logic of the Moral Sciences*. La Salle: Open Court

Popper, K. R (1966) *The Open Society And Its Enemies*. Princeton: Princeton University Press.

Sen, Amartya (1977) "Rational fools: A critique of the behavioral foundations of economic theory", *Philosophy and Public Affairs*, 6(4): 317–344.

Sperber, Dan (2011)"A naturalistic ontology for mechanistic explanations in the social sciences", in Pierre Demeulenaere (ed.), *Analytical Sociology and Social Mechanisms*. Cambridge: Cambridge University Press, pp. 64–77.

Turner, Jonathan H. (2003) *Human Institutions. A Theory of Societal Evolution*. Rowman and Littlefield Publishers.

Turner, Jonathan H. and Maryanski, Alexandra (2005) *Incest. Origins of the Taboo*. London: Paradigm Publishers.

Udéhn, Lars (2001) *Methodological Individualism: Background, History and Meaning*. London: Routledge.

Watkins, J.W.N. (1957) "Historical Explanation in the Social Sciences", *British Journal for the Philosophy of Science,* 8: 104–117. Reprinted in P. Gardiner (ed.) (1959) *Theories of History*. New York: The Free Press.

Weber, Max [1922] (1978) *Economy and Society*. Berkeley: University of California Press.

Wippler, R. (1978) "The structural-individualistic approach in Dutch sociology", *The Netherlands Journal of Sociology*, 14: 135–155.

Ylikoski, Petri (2012) "Micro, macro and mechanisms" in Harold Kincaid (ed.), *The Oxford Handbook of Philosophy of Social Science*. Oxford: Oxford University Press, 21–45.

7

MACRO MECHANISMS

Is Sociological Explanation Possible in the Absence of Action Theory?

Loris Di Giammaria
Sapienza University of Rome

INTRODUCTION

This paper considers the positions of Mohamed Cherkaoui and Peter Hedström on the issue of whether sociological explanation is possible at a macro level of analysis. In particular, the paper argues that Cherkaoui exhibits a certain flexibility in regard to the possibility that macro phenomena can be explained by mechanisms which are also located at a macro level. By contrast, Hedström seems to deny this possibility by maintaining that every explanation must be based on beliefs, desires, and interactions referable to individuals. Hence, the macro dimension consists of a set of structural parameters within which to frame individual action. From Hedström's perspective, which may be termed 'structural individualism' (see Hedström and Bearman, 2009), sociological explanation must therefore refer to the structural conditions of action; but this stance excludes the possibility of locating the explanatory mechanisms at the level of the social structure.

The position that Cherkaoui sets out particularly in *Invisible Codes* (2005) seems instead to opt for a model of explanation which permits, if necessary, reference to mechanisms of macro type which therefore do not include individual beliefs. What Cherkaoui seems to argue is that researchers do not necessarily possess knowledge such that they can refer to micro-level mechanisms. Because the conventional micro/macro distinction can therefore be drawn at a methodological and not ontological level, it is possible to envisage mechanisms that do not necessarily include reference to individuals and to a theory of action.

1. THE DEFINITION OF MICRO AND MACRO LEVELS: METHODOLOGY AND ONTOLOGY

The hypothesis presented in this section, which appears to span the positions of Mohamed Cherkaoui and Peter Hedström, expresses a difference of approach to the definition of micro and macro levels in explanatory sociological analysis. This difference seems to stand in direct relation to the definition and use of the notion of generative mechanism.

In particular, the conception of the micro–macro relations attributable to Cherkaoui seems to view these issues from a methodological perspective, thus going beyond the specific individualistic perspective that apparently characterises the views of Hedström and which locate explanatory mechanisms at an exclusively individual level or at that of relations between individuals. When, for example, Hedström refers to the approach of Blau and Black, he states that "a basic flaw with this mode of theorizing, as I see it, is that it excludes from the explanation exactly those processes that would have allowed us to understand why social entities exhibit the regularities they do. What are lacking in the approach of Blau, Black and other methodological "holists" are the basic entities and activities that generate these correlations" (Hedström, 2005: 19).

Unlike Hedström, Cherkaoui appears to recognise the possibility of reconstructing generative mechanisms located at the level of the social structure. He states that "there exist at least two generative levels; one deriving from mechanisms that do not have a micro explication and

one deriving from mechanisms with an explication of hypotheses on individuals" (Cherkaoui, 2005: 89). According to Cherkaoui, the necessity to explain a phenomenon by breaking it down into its individual or inter-individual components cannot be absolute. The need to identify explanatory hypotheses located at a micro or macro level depends on the type of problem that the researcher is seeking to solve. In this sense the approach delineated by Cherkaoui seems to set aside any hypostatization of ontological type, such as an imposition of the micro foundation of social phenomena in individual intentions. On the other hand, in regard to the ontological character of the necessity to found the explanation of social phenomena on individual action, Hedström is explicit: "The most reasonable ontological hypothesis we can formulate in order to make sense of the social world as we know it is that it is individuals in interaction with others that generate the social regularities we observe" (Hedström, 2005: 19).

By contrast, the position of Cherkaoui, which seems, as said, to support the methodological matrix of a possible reference to individual and intra-individual mechanisms, is set out in the following extract:

> if the paradigm of methodological individualism is necessarily associated with the strategy of generative mechanisms, it is not however always easy to apply. In fact, *because work in a particular area of research may not be well advanced, it is sometimes difficult or impossible for the time being to bring explanation down to the level of the individual.* Two options, then, are possible. The first, more radical, rejects as explanatory variables macrosocial phenomena or normative concepts that have not yet been explained and reduced to their individual components. *The second option is more flexible and will provisionally accept them* (Cherkaoui, 2005: 106, emphasis added)

And again, Cherkaoui states in his introduction to *Invisible Codes*, "I show that the most common definition of the micro and the macro, based on size or scale, is partial and even faulty. From this I suggest substituting a tridimensional definition that takes into account, initially and before anything else, the nature of the problem to be resolved, then

the type of hypothesis or of simplification that the researcher adopts to resolve it, and finally the nature of the observational unit and of the analysis that must be congruent with the first dimension" (*ibid.: 7*).

In *Dissecting the Social*, Hedström states that when a complex whole is resolved into its component parts, it should be the specific characteristics of the research problem that indicate the type of entity and activity to be regarded as central (see Hedström, 2005: 6). This statement suggests that, in sociological research, the entities and activities to be considered as constituting a complex system may be identified in both individuals and intra-individual components such as beliefs or desires, and thus (to use Coleman's words) with units that are not necessarily individual and at a level lower than that of the system (Coleman, 1990). For example Coleman refers to institutions or sub-groups that constitute the system, although he maintains that the most typical case is that in which the components are individuals who make up the social system. According to Hedström, it seems incontestable that "in sociological inquiries, however, the core entity always tends to be the actors in the social system being analysed, and the core activity tends to be the actions of these actors" (Hedström, 2005: 5).

Given these premises, it is appropriate to consider how the distinction drawn by Cherkaoui between a weak and strong version of methodological individualism seems to enable him to straightforwardly describe his proposal as a weak form of individualism, while Hedström's position seems more complicated. According to Cherkaoui, "for the strong version all explanation can only be conducted in terms of intentions, of motives or of the reasons of individuals without ever taking aggregated or emergent social phenomena". This in particular is the position adopted by Harsany (1968) and Hechter (1983). By contrast, "for the weak version social or macrosocial phenomena can be accepted as explanatory variables in the expectation that sociological theory can be better equipped to be able to deduce individual behaviour" (Cherkaoui, 2005: 54).

This latter version of individualism, in conformity with Popper's view, considers individual action to be necessarily influenced by social institutions. It maintains that the latter cannot be explained without reference to other institutions. The question evidently once again concerns

the location of explanatory mechanisms at an exclusively individual level or also at the level of the social structure. It recalls the solution proposed by Hedström, which in numerous contributions he has termed *structural individualism* (see Hedström, 2005: 5; Hedström and Bearman, 2009a: 8). According to Hedström, sociological explanations cannot be indifferent to already-existing social structures, even though, in order to achieve an adequate explanation, it is necessary to identify the micro mechanisms that at an individual level produce specific micro relations.

In the introductory essay to *The Oxford Handbook of Analytical Sociology*, which he co-authored with Bearman, Hedström clarifies the relation between individual intentions and social phenomena by presenting two models of the micro–macro relationship. The first model describes the relation between macro-type processes at two different times whereby both at t1 and at t2 certain individual actions express specific phenomena emerging at social level in the absence of a relation between the macro level at t1 and the micro level at t2. Otherwise stated, Hedström's first model does not envisage an impact by the social structure on individual actions; on the contrary, individual actions at t1 and t2 respectively express emergent social processes, in the same way that, in the body/mind relation, certain cerebral processes express successive emergent states in the mind (Hedström and Bearman, 2009a: 10). The

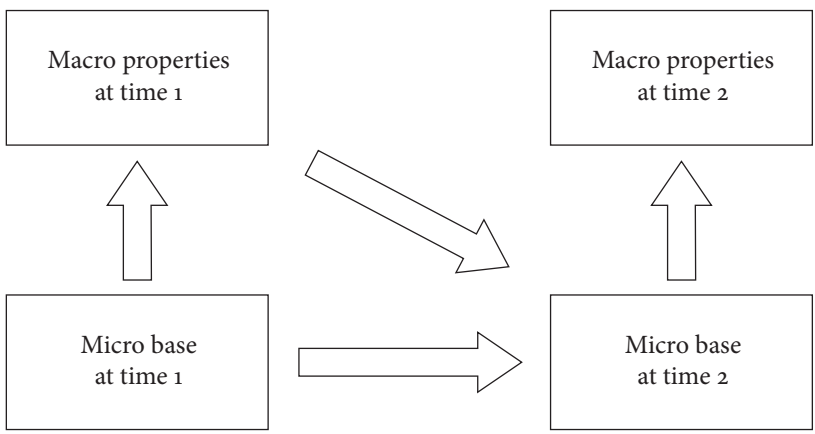


Figure 1: *The micro–macro relation according to structural individualism*

second model instead focuses on the nature of the approach that the authors name 'structural individualism'.

What emerges from Hedström and Bearman's model is the fact that the social structure exerts a conditioning impact on the micro/individual level at t2. According to the authors, this means that "macro-level properties also can have causal efficacy by being components which individuals consciously take into account when deciding what to do" (Hedström and Bearman, 2009a: 11).

Besides recognition of the role of the social structure in determining social phenomena, what emerges clearly is the necessity to locate explanatory mechanisms at a rigorously micro level, where 'micro' is understood as the level of individual action. The authors maintain that "the important point is that at each instance of time, macro M, is supervenient on micro, P, and in order to explain changes in M we must explain changes in this micro base" (*ibid*.: 12). For Hedström, every causally significant process attributable to the social structure must necessarily pass through identification of the mechanisms involved in individual actions/interactions. Hedström expresses himself in like manner in his previous work *Dissecting the social*, where he states that "although opportunities exist independently of an actor's beliefs, they must be known to the actor and hence they can be said to influence actions via the beliefs of the actor" (Hedström, 2005: 39).

From Cherkaoui's perspective, it appears possible to propound a third version of the model of the micro–macro relation, where the significant difference with respect to Hedström and Bearman's second model consists in the possibility to identify explanatory mechanisms at a macro level, and hence a causally significant relation at a macro-macro level. The structure of the model coincides with that of the 'Coleman Boat' (see Coleman, 1990). The difference from the version of the model attributable to Cherkaoui consists in the different emphasis placed on the macro-macro relation. Only in cases where the evolutionary phase of the research allows it, and where the selection of the cognitive objectives suggest the appropriateness of such an approach, is possible, according to Cherkaoui, to identify the individual and inter-individual mechanisms helpful for explanation of the social phenomenon under examination.

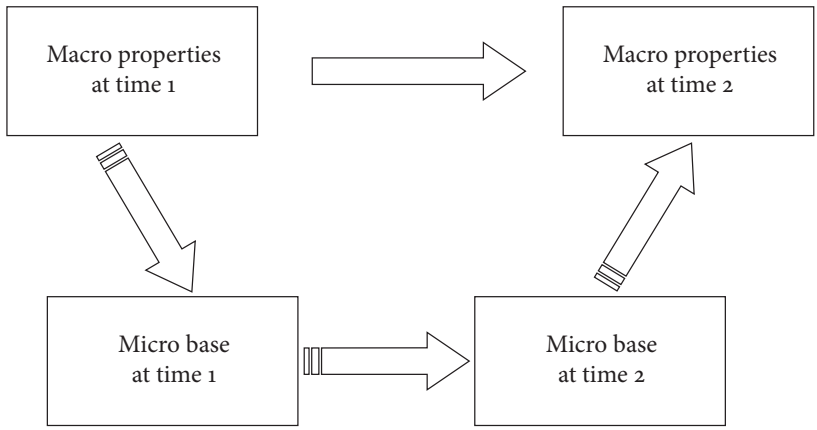


Figure 2: *The micro–macro relation in the Coleman Boat and in Cherkaoui*

According to Cherkaoui, in cases where the cognitive objectives of the research do not entail the need to postulate mechanisms at a micro-analysis level, one should not exclude the possibility of reconstructing the processes of interaction that may have brought about specific configurations of the social structure. This generates the following model.

According to Cherkaoui, "it is, of course, possible that the structures were themselves engendered by interaction processes (micrology). In this case the understanding would be that such processes brought

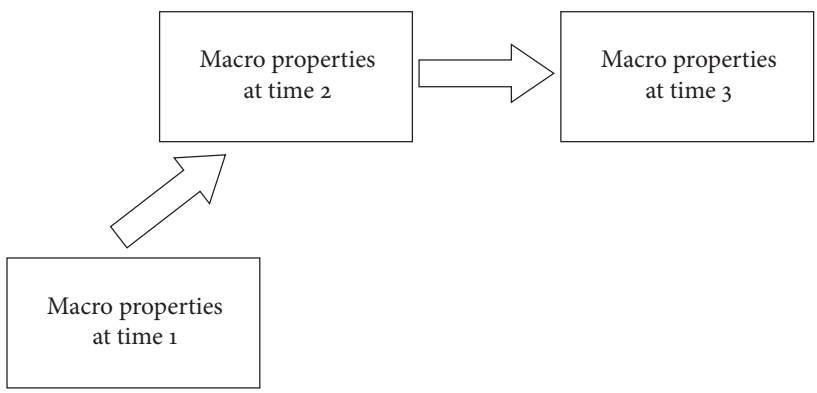


Figure 3: *The micro–macro relation in the formation of generative mechanisms at macro level*

about a differentiation of roles and positions which then became crystallized. It therefore differs from microsociology, which seeks to study the effect of preference on social relation" (Cherkaoui, 2005: 139).

This appears to mean that the interaction processes, once crystallized, assume an autonomous generative capacity. It is consequently possible, according to Cherkaoui, to reconstruct the genesis of particular structural elements, starting from certain processes of action and interaction located at to. This does not mean that once such elements have crystallized, they may not exercise a causal power, even in extreme cases entirely unbeknown to the individuals concerned. In Cherkaoui's opinion, one should consider the possibility that information on the configuration of position systems in certain sections of the social structure may be somewhat partial, so that it may happen that the effects of the configuration of the social structure affect individuals independently of their intentions.

2. GENERATIVE MACRO-LEVEL MECHANISMS AND EPISTEMOLOGICAL PLURALISM

One may gain more thorough understanding of what Cherkaoui means by 'generative macro-level mechanism' if we consider an example that he takes from Simon on the population densities attributable to certain cities. Let us suppose, says Cherkaoui, that we make a classification in ascending order in such a way that the population of each city is inversely proportional to its position in the classification. By creating a system of Cartesian axes and placing the position occupied in the classification on the abscissa, and the number of inhabitants on the ordinate, one sees that the concentration of the points representing the cities describes a downward sloping line approximately equivalent to –1. According to Cherkaoui, "the mathematical law thus revealed describes but does not explain. A generative mechanism must be constructed. This mechanism logically takes into consideration the necessary demographic elements which, when combined according to the relations to be defined, must produce the empirical results and the mathematical law describing them" (Cherkaoui, 2005: 50).

According to Cherkaoui, one needs to construct a generative mechanism that combines, on the basis of relations to be defined, the

mortality and birth rates, migratory flows and the appearance of new cities. On considering a territory that includes both urban and rural areas, Cherkaoui continues, it is possible to hypothesize an absence of correlation between birth and mortality rates and the sizes of the cities, and between rates of emigration from the cities and their sizes; while it seems possible to hypothesize a correlation between rates of migration to the cities and their sizes. Furthermore, he continues, it is possible to hypothesize that the migration from one city to another is regular and quite distinct from migration from rural areas to the cities (see *ibid*., 2005: 50).

A further example of how it is possible to use macro-level mechanisms is provided by Cherkaoui with reference to the relation between religious faith and the level of mixed marriages between religious groups. Having ascertained that the percentage of Protestants that marry non-Protestants is higher than that of Catholics who marry non-Catholics, it is possible to explain these significant differences only by considering the distribution of individuals in certain positions and the effects of such positions on the individuals. Let us take, Cherkaoui says, two groups of different sizes, with A<B; the individuals of group A are able to establish relationships with individuals of group B and marry them. Examination of the respective rates of mixed marriages, he concludes, allows us to say that the rate of mixed marriages is higher in group B than in group A, and that the correlation initially considered can be explained in terms of the sizes of the groups (see Cherkaoui, 2005: 140).

For Cherkaoui, macro-level mechanisms can be of three types:

1) Mechanisms such as systems of hypotheses concerning relations between macro variables (relative to aggregate behaviour) that act on a specific *explanandum* phenomenon, e.g. the size of a city in the above example.
2) Mechanisms identifiable with the structural properties of more or less formalized collective units (groups, institutions, etc.); such properties act on the distribution of certain positions in the social structure, as in the example of the impact of the number of religious groups on the marriage market. The abstract forms of social cohesion, Cherkaoui states by citing Simmel, are "materializations of group *forces* and they therefore have a supra-individual, objective

character that stems from the multitude of their elements and their effects, with which the individual is confronted"(Simmel, 1908, from Cherkaoui, 2005: 139, our italics).

3) Mechanisms identifiable with the action/interaction between structural components of a level lower than the structural unit originally identified, and which in turn influence the distribution of positions within the social structure. The reference here is to analysis of the social stratification and the relation between educational and income levels which considers the positions in the education system and labour market to be fundamental explanatory mechanisms (see Cherkaoui, 2005: 145–160).

If we consider the notion of mechanism used by Hedström, it is apparent that Cherkaoui's notion has a broader frame of reference and includes, not only the entities, but their respective properties and the activities carried out by them (see Hedström, 2005: 32); a single property can be considered a mechanism (for example the size of a group), in the same way as a system of hypotheses can be referable to variables. In fact, the notion of mechanism used by Cherkaoui suggests that there are manifold ways to explain certain social phenomena, and that any explanatory account that clarifies an original relation may be considered a mechanism,. A black box is thus opened. In this regard, Cherkaoui seems to agree that it is necessary to open black boxes in order to effectively advance explanatory accounts; but at the same time he appears to suggest 1) that every black box may lead to a further black box that needs to be opened; 2) that in certain research problems there is no need to open certain black boxes (located, for example, at an individual or intra-individual level) because they are irrelevant.

The possibility to use multiple types of generative mechanisms that are not necessarily comprised in a theory of action, and thus do not necessarily have individuals as points of reference, would seem to be the result of an analytical approach directed towards a theoretical-epistemological pluralism. Cherkaoui expresses himself very clearly on this matter: "it should be apparent that the macro structuralist project is clearly circumscribed, limited to certain aspects of the real. While rational choice theory presents an undeniable advantage over other theories, it

cannot serve as a universal panacea; presenting it as unconditionally valid makes it vulnerable to the same dangers that other theories have encountered. As for functionalism, its fatal error was to succumb to the temptation of hegemony and imperialism; it claimed the title of general theory of social systems, when some of its principles are only valid for particular, tightly circumscribed domains" (Cherkaoui, 2005: 141). The typology of theories drawn up by Cherkaoui, in fact, goes in this direction. It distinguishes social theories on the basis of three parameters: 1) the nature of the specific problems being studied (for example, those of the regularity of individual behaviour and social or group regularities). In this regard, Cherkaoui stresses that one must pay attention to the fact that some problems that appear to be positioned at a macro level – for instance, the effects of the occupational statuses of fathers on the social positions of their children – should in fact be considered at a micro level of analysis; 2) the type of hypotheses, rational or otherwise, used to furnish the explanation; 3) the type of unit being observed, such as individuals, situations or structures.

In the name of theoretical pluralism it would thus seem possible to use as explanatory mechanisms entities identified with social forces that result from a type of crystallized structure referable to certain collective groups.

Moreover, it seems that also Merton identified mechanisms operating at the level of the social structure, in the form, for instance, of social mechanisms that articulate roles within the role complex, or the mechanism relative to power differences among fringe positions within the role complex, etc. (see Merton, 1968: 426–33).

The forces located in the social structure – consistently with an approach referable, for instance, to critical realism – may be considered *powers intrinsic* to a certain type of collective unit of analysis. That is to say, the presence in a particular collective group of specific structural characteristics may imply the presence of certain dispositions of that collective group. It thus seems that Cherkaoui's position encompasses the following aspects:

1) Postulation of a social and individual reality that is independent of the conceptual activity of the cognitive subject.

2) The attribution to a micro and macro level of distinct entities that are causally significant (this does not mean that the contents of the two levels are definable once and for all).
3) The identification of the entities, of both micro and macro type, with specific powers, or dispositions.[1] It thus seems that the causally significant entities are identifiable with the *specific potentialities* developed, for instance, by certain institutions, rather than with the institutions as such.

Hedström does not seem to concur on these points. Analytical sociology, in fact, seems deny that it is possible to identify the micro and the macro with levels of reality that are ontologically distinct, in that the macro is no more than the more or less conscious expression of the micro. This is not to attribute a reductive position to Hedström, but rather the idea that a correct explanation requires the micro-founding of every emergent social phenomenon.

Hedström's attack against critical realism is harsh. His accusation is that critical realism envisages causal powers of a rather mysterious nature, which seem to expand into a decidedly metaphysical universe. These are mechanisms, Hedström concludes, which are not anchored to a valid methodological proposal (Hedström, 2005: 72–3). This is, in effect, a criticism that can be countered if one argues in two distinct directions: 1) as Archer clearly states, the notion of ontology referred to by the critical realists, is not of an essentialist type; 2) the powers referred to by the critical realists can be conceptualised as dispositions, and the concepts that denote such powers as dispositional concepts.[2]

The concept of reality that Cherkaoui uses particularly in the chapter of *Invisible Codes* entitled "Multiple Levels of Reality" (see Cherkaoui, 2005: 115) suggests an ontological dimension in the distinctions between the levels. The micro and the macro are defined as *scales of organisation*, in such a way as to refer 1) to the size of the units of analysis, where 'size' is understood as the breadth of the relational configurations referred to (from the single individual to complex institutionalised systems); 2) to the type of organization, whether related to a configuration of individual relations or to collective groups characterised by different levels of institutionalisation.

In this regard, it is of interest that Cherkaoui does not refer to two levels of reality, but *to multiple levels of reality*, in order to underline that the micro and macro are not definable once and for all, and that the levels of reality in the social sciences may multiply, as has occurred in physics, and express specific levels of stability and latency (see Cherkaoui: 117). Cherkaoui acknowledges that there has not been a clear codification of the notion of stability or of individuality in the social sciences (see *ibid*.: 107), yet he implies the crucial importance of these concepts. Here one might ask *whether the stability referred to is related to certain outputs from the mechanisms, or to the mechanisms themselves*. In the first acceptation of stability one may imagine a succession of progressively different mechanisms, each of which produces certain outputs with a certain regularity; in the second acceptation one may imagine the action of the mechanisms themselves as they persistently produce certain outputs with a certain regularity. The epistemological pluralism of Cherkaoui in the sense just described seems to encompass a multiplicity of possible mechanisms characterised by different levels of stability and latency.

In analytical sociology, even in the context of a declared realism, the possibility/necessity to attribute ontological consistency to the macro level is denied because the only entities considered to really exist and act causally *directly at the macro level* are individuals and their respective mental states[3]. Accordingly, Hedström's realism seems to have more to do with a form of epistemological phenomenalism that seeks to reconstruct configurations of unstable phenomena easily accessible to observation (the intentions and desires of individuals and their interactions) than with forces characterised by significant latency or persistency (dispositions/potentialities attributable to both single individuals and collective groups) (see Di Giammaria, 2010). It is therefore possible to speak of phenomenalism because the denial of the causal significance of a multiplicity of possible mechanisms, except those relating to individuals and their interactions, seems to rule out the possibility of a realism based on some idea of the stability/persistence of the mechanisms themselves (see Harrè and Secord, 1972). Hedström's position appears to coincide with a sophisticated version of constructivism which maintains that scientific theories must produce descriptive and factual

statements irrespective of whether they refer to levels of profound and persistent reality.

3. CONCLUSIONS: SOME CONSIDERATIONS ON GENERATIVE MECHANISMS AND STRATIFICATION OF REALITY

On the basis of what has been discussed it is possible to indicate some matters for final reflection. The first are specifically related to the notion of generative mechanism, while the second concern the notion of sociological theory and a description of the micro–macro relation in regard to a stratification of levels of reality.

1. Continuity in the notion of mechanism in Cherkaoui and Hedström
The definition of mechanism proposed by Hedström, which is at a high level of abstraction, can include reference to mechanisms of an entirely social nature, whose constituent entities undertake action that is causally significant for a specific *explanandum* phenomenon. Both authors consider the explanatory function of the macro-level to be important. It is necessary to conceive the macro level as the set of structural constraints imposed on individual action (see Hedström and Swedberg, 1998). This level of analysis is present for example in the first empirical works of Hedström (see Hedström, 1992, 1994). On this point, the difference between the positions of Hedström and Cherkaoui seems blurred, or even disappears, except for what appears to be a significant distinction expressed in the following point.

2. Discontinuity in the notion of mechanism in Cherkaoui and Hedström
Compared with Hedström's position, that of Cherkaoui on the notion of mechanism seems to diverge: where Hedström considers the individual as the constitutional and indispensable entity for sociological analysis, Cherkaoui seems, instead, to take a pluralistic position on which entities are to be regarded as explicatively relevant, so that it *includes reference to certain types of entities related to specific phases of the research and specific types of cognitive problems.*

In other words, whilst Hedström considers reference to the macro conditions of individual action to be useful, Cherkaoui explores whether it is possible to identify also at macro level (and as a limiting case only at the macro level) the explanatory mechanisms with which to achieve particular knowledge goals, given a certain level of development of the research process. For Cherkaoui, the macro-macro relation can acquire a causal thickness: the socio-cultural structure does not necessarily have to operate through the micro level to influence macro phenomena.

Cherkaoui's position therefore seems somewhat open to the idea that the entities to be considered causally active at the macro level are not only individuals but also structures, possibly endowed with autonomous causal powers.

In this sense, the notion of mechanism considered by Cherkaoui seems consistent with the general notion of mechanism propounded by Hedström (Hedström, 2005: 25). But it may have a semantic reference broader than the specific interpretation given to it by Hedström (see *ibid*.: 26; Hedström, Ylikoski, 2014: 64). The latter believes that although rules and institutions can play a significant role in explanation, in particular in the macro-micro relation (Hedström, Ylikoski, 2014: 64), "the aim of explanation is to identify those beliefs, desires, or goals that make the causal difference to an individual's behavior" (Hedström and Ylikoski, 2014: 64). For Hedström, beliefs and desires are not necessarily related to intentionality (see *ibid*.: 65). He therefore argues against

Table 1: *Types of mechanism at the macro level*

Definition	References
Systems of hypotheses about variables that denote aggregate behaviours	*Compositional Effects and Generative Mechanisms*, in Cherkaoui (2005: 50)
Properties attributable to collective units of analysis	*The Multiple Levels of Reality*, in Cherkaoui (2005: 140)
Processes of interaction between lower-level structural components and the structural unit originally identified	*Notes on the Syntaxes of Stratification Theories*, in Cherkaoui (2005: 145–60)

the absolutism of the theory of rational choice and against intentional fundamentalism (Ylikoski, 2012: 39).

Table 1 summarizes the types of mechanism that can be positioned at the macro level according to Cherkaoui.

Furthermore, as regards other implications, it is possible to develop the following considerations by theme:

a. The relation between mechanisms and multi-variate analysis. With particular reference to the example given above on urban population density, Cherkaoui's explanatory hypothesis seems amenable to multivariate analysis where the hypothetical model (identifiable with the generative mechanism) can be tested by empirical correlation. In other words, one may say that the mechanism in the case of the example is constituted by the set of explanatory hypotheses referable to relations between variables to be empirically tested. These are hypotheses that specify the original relation between the positions in the classification of the cities and their population densities by referring to birth and mortality rates and migratory flows, and they represent an explanatory model that can be operationalized.

Hedström seems willing to recognize that the establishment of links between agent-based models and quantitative sociological research can bring benefits to both traditions (see Hedström, 2005: 111). What he does not seem willing to admit is the possibility of identifying the generative mechanisms with hypotheses on relationships between variables.

b. Plurality of the entities and levels of analysis. Cherkaoui's reasoning seems consistent with the idea that one cannot take the possibility to open black boxes as absolutely certain for the simple reason that, once a black box has been opened, it will immediately be possible to hypothesise another black box located at a deeper level of latency. Cherkaoui thus seems to address the issue of the plurality of explanatory levels, and he does not exclude the possibility of explanatory accounts based on levels of analysis not identifiable with those of individuals or relations between individuals. His position appears to exclude any hypostatisation of the type of entity or level of analysis within which to locate the black boxes to be opened. This makes it possible to hypothesise that

individual beliefs and desires do not of themselves constitute the entities that drive social action and processes, and that it is possible to postulate the action of different elementary entities, both in individuals and collective groups, that are characterised by an even higher level of latency.

a) Individuals: dispositions relating to a structure of the individual's personality, or even to the neurological structure of the brain (see Barbera, 2006b: 156; Di Giammaria, 2009: 170–92).

b) Collective groups: dispositions attributable to components of the social structure that are themselves structural (groups, institutions etc.). In this regard, Cherkaoui hopes for and predicts a codification of macro-sociology and its elementary units in keeping with the perspective indicated by Lazarsfeld (see Cherkaoui, 2005: 160).

c. The role of sociological theory in the selection of mechanisms and explanatory levels. The plurality of the entities acting as generative mechanisms and the plurality of explanatory levels that we have considered suggest that Cherkaoui adopts a position of epistemological and theoretical pluralism. Put otherwise, decisive for the selection of the explanatory levels and the types of mechanism to be considered as acting on certain social phenomena, is the theory to which one refers. The latter may be appropriate both to specific cognitive objectives singled out by the researcher, and to specific levels of *theoretical maturity* – as Cherkaoui states consistently with Hedström's position.

d. Stratification of reality and ontology. From the description of the types of mechanism considered by Cherkaoui it is possible to draw some conclusions on his conception of the micro–macro relation. *Cherkaoui appears to propose a theoretical approach to the question.* This means that he seems to bear in mind that in sociology the definition of what is micro and what is macro cannot be given once and for all; it is in fact an attribution that always stands in strict relation to the research problem under examination (in an investigation on voting behaviour in a specific country, the macro level may be identified as the aggregate vote in each province, and the micro level as the aggregate vote by constituency; in an investigation on the selection of seats in a cinema, the macro level is given by the final seating arrangement in the cinema, and the micro

level by the reasons for which the single members of the audience select specific seats). Consequently, Cherkaoui seems to identify the distinction between micro and macro as that between the levels of reality, in the sense that the two levels premise a reference to distinct entities at a substantive level. It is therefore possible to think in terms of different ontological levels, but on one specific and significant condition: it is necessary to consider an ontology stripped of its essentialist dimension and rigorously transferred within the confines of empirical scientific debate. Here it is necessary to clarify an apparent contradiction in Cherkaoui's work, and which relates to his description of micro and macro levels not as mere levels of analysis but as levels of analysis that correspond to a reality while at the same time negating their ontological consistency (see *ibid*.: 116); how is it possible to reconcile the negation of an ontological reference with the claimed use of the concept of reality? The answer appears to reside in the distinction between two specific meanings of the notion of ontology: a concept of ontology that Cherkaoui clearly wishes to be rid of, and an acceptation that he appears instead prepared to defend. The concept of reality used by Cherkaoui appears coherent with that applicable to physics, in an acceptation of reality that is distinct from essentialist metaphysics but coherent with the description of entities and forces existing at a real level.[4]

The difference from analytical sociology in regard to this question can be attributed to the fact that Hedström, by hypostatising the individual's explanatory role without any possible alternatives, raises suspicions of a covert dogmatic essentialism. The explanatory importance given to structures must in fact necessarily pass through individual beliefs and desires. This rigidity can perhaps be explained as a reaction to the array of possible and devastating essentialisms to be found in the holistic domain, while, paradoxically, uncovering some 'germs' and reflections of the same. From a genuinely scientific perspective, there persists the necessity (which would seem to be one of Cherkaoui's key reference points), to let the results of the research, and their usefulness for the evolution and epistemic strengthening of the social sciences, decide which entities are to be recognised as generative mechanisms.

NOTES

1. The notions of *disposition* and that of *dispositional concept* have been subject to debate in both logic and the philosophy of science, as well as in the discussion addressed strictly to the social sciences. Lazarsfeld, in his essay on the *formation and measurement of concepts in the behavioural sciences* (1966), uses the notion of disposition superimposing it on the notion of *trait*. Dispositions, which are definable as *stable tendencies* of certain objects that manifest specific reactions in specific circumstances, in the social sciences seem attributable for Lazarsfeld to both individuals and collective groups. It is thus a matter of hypothesizing the existence of certain *intrinsic powers*, to use a concept attributed, for example, by Goodman (1954) to the notion of disposition, which would appear akin to the notion of causal power used by critical realism.
2. Various proposals for the operationalization of dispositional concepts in the social sciences have been put forward to solve the problem of the limited methodological anchorage of conceptual models that postulate the causal action of dispositions (see Di Giammaria, 2009).
3. When reference is made to the notion of causality, the intention is to use the idea of causality that the analytics consider interconnected with the notion of generative mechanism, and in particular with the general definition of mechanism put forward by Hedström (Hedström, 2005: 25). The applicability of this notion to the macro-macro level is a possibility that Cherkaoui apparently envisages when he proposes the idea of macro mechanisms that account for empirical uniformities in their turn of macro type.
4. It is possible to relate the idea of a reality as existing independently of linguistic or conceptual activity to an example proposed by Quine. For this author, the capacity to translate a question into semantic terms does not mean that it is of itself a linguistic question. For instance, "to see Naples" entails using a name that forms a real enunciation, and there is nothing linguistic in the fact of seeing Naples (see. Quine, 1963: 16). Support for a sophisticated realist position is given by many authors of the theory of science, ranging from Khun to Hacking and Feyerabend. The position on this question, which the last of these authors attributes to Einstein, consists of a scientific realism presented as a working hypothesis and not as a conception of the world (see Feyerabend, 1999).

REFERENCES

Barbera, F. (2006) "Meccanismi generativi, teoria dell'azione e struttura sociale (M. Cherkaoui, *Invisible codes*, Oxford, The Bardwell Press (2005)", *Sociologia e Ricerca Sociale*, 80.

Cherkaoui, M. (1979) *Les Paradoxes de la réussite scolaire. Sociologie comparée des systems d'enseignement*. Paris: Presses Universitaires de France.

——— (2002) *Macrosociology/microsociology*, in Smelser (2002).

——— (2003) "Stratification sociale et niveaux de réalité", in Cherkaoui (2003).

——— (ed.) (2003) *Histoire et théorie des sciences sociales: mélanges en l'honneur de Giovanni Busino*. Geneva: Libraire Droz.
——— (2005) *Invisible Codes: Essays on Generative Mechanisms*. Oxford, The Bardwell Press.
Coleman, J. (1990) *Foundation of Social Theory*. Cambridge, MA: Harvard University Press.
Di Giammaria, L. (2009) *Realismo scientifico e disposizioni sociali. Teoria sociologica e metodologia della ricerca sociale nel dibattito tra realisti e costruttivisti*. Catania–Rome: Bonanno.
——— (2010) "Concetti di meccanismo. Stabilità, latenza, e alcune implicazioni teorico-metodologiche nelle scienze sociali", *Sociologia e Ricerca Sociale*, 91.
Di Renzo, G.J. (1967) *Concepts Theory and Explanation in the Behavioral Sciences*. New York: Random House.
Feyerabend, P. (1999) *Conquest of Abundance*. Chicago: The University of Chicago Press.
Goodman, N. (1954) *Fact, Fiction and Forecast*. London: Athalone Press.
Harrè, R., Secord, P.F. (1972) *The Explanation of Social Behavior*. Oxford: Basil Blackwell.
Hedström, P. (1992) "Organizational vacancy chains and the attainment process", *Journal of Mathematical Sociology*, 17: 63–76.
——— (1994) "Contagious collectivities—On the spatial diffusion of Swedish trade unions, 1890–1940", in *American Journal of Sociology*, 99: 1157–1179.
——— (1998) "Rational imitation", in Hedström and Swedberg (1998b).
——— (2005) *Dissecting the Social. On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.
Hedström, P. and Bearman, P. (2009a) "What is analytical sociology all about? An introductory essay", in Hedström and Bearman (eds) (2009b).
——— (eds) (2009b) *The Oxford Handbook of Analytical Sociology*, Oxford, Oxford University Press.
Hedström, P. and Swedberg, R. (1998a) "Social mechanisms: An introductory essay", in Hedström and Swedberg (1998b).
——— (eds) (1998b) *Social Mechanisms. An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.
Hedström, P. and Udehn, L. (2009), "Analytical sociology and theories of the middle range", in Hedström and Bearman (eds) (2009).
Hedström, P. and Ylikoski, P. (2014), "Analytical sociology and rational-choice theory", in G. Manzo (ed.), *Analytical Sociology. Actions and Networks*. Chichester: John Wiley & Sons.
Kincaid, H. (ed.) (2012) *The Oxford Handbook of Philosophy of Social Sciences*. Oxford: Oxford University Press.
Lazarsfeld, P.F. (1966) "Concept formation and measurement in the behavioral sciences: Some historical observations", in Di Renzo (1966).
Quine, W.O. (1963) *From a Logical Point of View. Logico-Philosophical Essays*. New York: Harper & Row.
Turner, B.S. (2005) "Foreword. Visible forms and invisible codes: Description, interpretation and explanation", in Cherkaoui (2005).

Ylikoski, P. (2012) "Micro, macro and mechanisms", in Kincaid (eds) (2012).

Additional Readings

Barbera, F. (2006), "Il ritorno del realismo analitico", in Hedström (2006).
——— (2008) *Realismo critico e (o?) realismo analitico*, in Maccarini, Morandi, Prandini (2008).
Borlandi, M. and Sciolla, L. (eds) (2005) *La spiegazione sociologica. Metodi, tendenze, problemi*. Milan: il Mulino.
Lombardo, C. (2012) "The missing heritage. On Lazarsfeld's contribution to the development of analytical sociology", *Sociologia e Ricerca Sociale*, 97.
Maccarini, A.M., Morandi, E. and Prandini, R. (2008) *Realismo Sociologico. La realtà non ama nascondersi*. Genoa–Milan, Marietti.
Manzo, G. (2010) "Analytical sociology and its critics", *European Journal of Sociology*, 51(1): 129–70.
——— (eds) (2014) *Analytical Sociology. Actions and Networks*. Chichester: John Wiley & Sons.

Not in references.

8

THE STRATEGY OF MECHANISM-BASED RESEARCH

Diego Ríos

Consejo Nacional de Investigaciones Científicas y Tecnológicas (CONICET), Buenos Aires, Argentina.

UNTREF, Buenos Aires, Argentina.

1. MECHANISMS, LAWS AND EXPLANATIONS

Mechanism-based explanations are of superlative importance both in the natural and the social sciences. They constitute a fruitful research strategy that has been applied with success to a wealth of phenomena—from cell biology (Bechtel, 2006) to the formation of racial ghettoes (Schelling, 1969) and the market of used cars (Akerlof, 1970). The crucial point that unifies the mechanistic-based account is the commitment to explain the high-level properties of a given system in terms of the behaviour of its components. However there are important differences in how these components generate the high-level behaviour of the system (Cherkaoui, 2005). In some cases these components are directly responsible for fulfilling specific, dedicated sub-tasks—as in the case of a mechanical clock, where each one of the components is assigned a relatively discrete role, different in any case to the roles fulfilled by other sub-components. In other cases, the components do not differ in the tasks they perform; they all do the same

thing, and the resulting properties of the system are the result of their interaction.

Despite its wide use in science, mechanism-based explanations only recently captured the attention of philosophers of science. The reason for this delay is probably due to the influence of the positivist orthodoxy that has dominated this discipline at least from the early 1930s. According to the received view, scientific theories were conceived as abstract calculi—as logical machineries capable of generating inferences. The main objective was to reconstruct current scientific theories in axiomatic terms. Laws were of paramount importance in this system. Laws constituted the axioms of the system and the observational statements were the theorems. This conception of scientific theories had a deep impact on the understanding of the tasks of science. In fact, in parallel to the syntactical conception of theories, positivist philosophers developed a syntactic account of explanation and of theory reduction (Nagel, 1961). According to this account, explanations are conceived as arguments having a specific inferential structure: in fact, explaining is deductively linking the axioms (*explananda*) with the theorems (*explanandum*). The key aspect of the explanatory process is subsuming observational statements under universal laws (Hempel, 1965).

This model of explanation is essentially epistemic: to explain an event is to show how this event would have been *expected* to happen. Explaining amounts to removing the surprise concerning the occurrence of a given event (Psillos, 2002). Laws play an essential role in the grounding of epistemic expectations. According to Hempel, a given phenomenon is explained when it is made the logical consequence of a general law and a set of initial conditions describing the antecedent conditions of application of a law. Suppose that we want to explain why a certain piece of matter—call it M—expands (*explanandum*). The explanation consists of stating a law—for instance, that metals expand when heated -, together with a set of conditions establishing that M is in fact a metal and that M is heated. If the premises of the *explananda* are true, we can nomically expect that M will expand.

This explanation does not mention the mechanism by which heating generates expansion. We know that heating causes metal to expand, but we do not know which mechanism underlies this process: we do not

know why heating causes expansion in metals. The explanation does not explicitly make the connection between the variables, it just states a law. Note however that the syntactic framework does not directly reject the notion of mechanism. It rather downplays its role. Explaining the behaviour of a system is equivalent to identifying the laws governing it, but without necessarily making the underlying causal processes involved in the generation of the output explicit.

Yet, an important part of the scientific enterprise consists of accurately identifying the actual causal process governing natural phenomena. This is the task which mechanisms are supposed to fulfil. Lacking a mechanism, the explanation will be a black box explanation (Bunge, 1964). Many theories have a black box structure—think for instance on behaviourist psychology. It seems, however, to be an important part of scientific explanation to go beyond the isolation of high-level regularities to try to identify the causal process—the wheels and cogs—of a given phenomenon. Take for example Weber's classical analysis of the raise of capitalism (Weber, [1905] 2009). He first identifies a correlation between Protestantism and capitalism: he observes that areas having higher rates of affiliation to Protestantism experienced a more stable and robust development of capitalism. Yet the most interesting part of Weber's analysis concerns the details making the connection between Protestantism and capitalism explicit—notably, the role of Calvinist ideas about predestination; these ideas generated strong anxiety amongst Protestants for finding signs of God's favour in economic success. The details of the mechanism explored by Weber are too rich to be captured by the traditional model of explanation.

The neglect of mechanism in traditional philosophy of science is not surprising. The introduction of mechanisms in the structure of scientific explanations makes these explanations more difficult to accommodate within an axiomatic model cherished by positivist philosophers of science. Laws have a relatively simple formal structure, and this formal structure can be smoothly captured with the apparatus of first order logic. The syntactic approach, without explicitly rejecting the role of mechanisms in science, provided an environment that was not congenial to it. This explains why the mechanistic account took so much time in attracting the attention of most philosophers of science. Things

started to change with the development of the semantic conception of theories (Giere, 1988; van Fraassen, 1980). Progressively, it became salient that most theories are recalcitrant to an axiomatic rendering. Furthermore, even when the axiomatic unpacking is possible, it is not obvious that it will always deliver the promised epistemic advantages. Although axiomatic systems are sometimes powerful theoretical tools, they are sometimes error-prone; the smaller the number of axioms in a theory, the more complex the structure of the deductive chain, and the easier it is to make inferential mistakes (Wimsatt, 2007). For this reason, axiomatic theories are seldom used for teaching and learning in science (Giere, 1988). The progressive development of the semantic conception of theories, with its emphasis on models rather than on theories, and its focus on representation rather than on logical deduction, thus provided a friendlier environment to theorize mechanistic explanations. It is then not surprising that the increasing attention given to mechanism-based explanations flourished concomitantly with a shift toward the semantic conception of theories.

2. THE STRUCTURE OF MECHANISTIC EXPLANATIONS

The thrust of mechanistic explanations is to account for the behaviour of a system in terms of the behaviour of its components. Mechanistic-based explanations are essentially analytical, in the sense that they attempt to disaggregate the phenomena under study, in order to later show how their components interact to produce the high level behaviour (Ríos, 2004). There are consequently two moments in the unfolding of a mechanistic explanation. The first step is the identification of the phenomena that needs to be explained. In most cases the phenomena can be unpacked as a causal regularity, one variable having an effect on another, and the challenge is accounting for the connection between these two variables (van Parijs, 1981).

The second step is to provide the mechanism. This might involve detecting the connection between the independent and the dependent variables, or simply, accounting for the behaviour of the system under study by describing how it can do what it does. An important constraint

in this second moment of mechanism-based explanations is that it must provide the rules of the system as autonomously as possible. The rules must be able to produce the effect without external intervention. This is the hard core commitment of the mechanistic account. The autonomy of the generative process is crucial to the mechanistic approach, and this autonomy must be obtained through a purely naturalistic process. Mechanistic explanations have been associated from the very beginning with the elimination of mysteries from the natural world. The model of the clock is again useful to fully grasp this point. A clock might require an external input to work—think about the sundial, water clocks or even mechanical clocks requiring external stimuli to be charged– but none of these exogenous factors must involve non-natural phenomena. The introduction of non-natural factors would inevitably threaten the mechanistic project.

It is interesting to note nevertheless that the autonomy required by mechanistic explanations does not mean that the mechanism must be as detailed as possible. The mechanistic model only requires that the mechanism must be sufficiently detailed to be able to promote the behaviour in question—to account for the behaviour of the system under scrutiny. Mechanistic explanations are sometimes understood as entailing a preference for fine-graineѕs—to wit, for a completely detailed account of the micro-process generating a given high-level phenomenon. This is however not right. Mechanistic explanations are not committed to the requirement of fine-graineѕs; they are rather agnostic about the degree of details that should be included in the specification of the micro-processes underlying the behaviour of the system. The important point is that the micro-processes should be able to account for the behaviour of the system. Craver argued convincingly that mechanistic explanations normally involve many levels—they are essentially inter-level explanations (Craver, 2007). In a recent publication we explored in mechanistic terms an intriguing phenomenon having an essentially inter-level nature (Kuchle and Ríos, 2012).

A crucial aspect of the mechanistic model is its accrued accuracy and reliability. The information provided by a mechanistic explanation is more reliable than the information provided by black box explanations. They provide richer insight about the phenomenon that we want

to understand, and this increased accuracy has important advantages. Knowing the causal process governing a phenomenon is essential to have a complete picture of the system to be accounted for. Furthermore, this epistemic advantage comes with an extra pragmatic benefit. If we are interested in policy intervention, having identified the mechanism generating a certain high-level effect is crucial for the success of our decisions to alter or modify the behaviour of the system. Black box explanations are poor guides for intervention.

An example from the history of medicine can illustrate this point. The causes of *diabetes mellitus* have been ignored for a long time (Bliss, 1984). Renaissance doctors already disposed of some description of the main symptoms, including the paradigmatic existence of sugar in urine from which this sickness takes its name (mellitus = honey). They obviously ignored both the organs involved as well as the biochemical mechanisms of the disease. They did not have the appropriate technology to measure the level of glucose in the serum either (in the bloodstream); they relied only on urine samples and they used taste to determine the existence of sugar. With this information, and ignoring the true mechanisms governing the illness, they thought that the sickness should be the result of a loss of sugar: patients with diabetes had extremely sweet urine. They consistently advised their patients to maintain a diet high in calories, increasing the intake of sugar in order to compensate for the loss of sugar in the urine! It took a long time for science to understand the deadly implications of this treatment. We know today that the loss of sugar in the urine is an automatic compensation mechanism for the excess of sugar in the blood. The mechanisms of the disease became clear at the beginning of the last century, with the great discoveries in endocrinology, notably with the discovery of insulin in 1921–1922 by two Canadian scientists (Bliss, 1984). Once this mechanism was in place, the whole picture changed. Diabetics have a malfunctioning pancreas that produces very low levels of insulin or no insulin at all, and consequently leads them to have an excess of glucose in the bloodstream. The existence of glucose is just a consequence, not a cause, of the sickness. Having identified the mechanism of the sickness has been crucial to provide the right treatment. Note that the original explanation of Renaissance doctors relied on information that was essentially correct—diabetic patients

have abnormal levels of sugar in their urine. Their reasoning was not implausible: it was reasonable to think that diabetic patients should be given extra sugar to compensate for the high-levels of sugar loss. The problem was not that Renaissance doctors used false information; the problem is that they made a wrong—although plausible—inference; their mistake was the consequence of not having a complete understanding of the mechanism involved in diabetes.

3. COMPOSITIONALITY

There are different ways in which these components interact. An important aspect to take into account is the degree of *exchangeability* of the components. In this section I would like to discuss decomposable systems—to wit, to wit, systems in which their components perform discrete, specialized subtasks (Simon, 1969; Betchel & Richardson, 1993). Suppose we want to explain the functioning of a car. We can decompose it in different parts, each one of them dedicated to carrying out a specific task. The engine, the fires, and the wheels are discrete components, and each one of them contributes to the functioning of the car in a specific way. The essential feature of this system is that the components are not fungible because they are task-specific—each one of these components is assigned a fixed, discrete role within the whole system, and it cannot be used for a different purpose. The fires cannot be used, for instance, to produce combustion.

The goal is to explain the behaviour of the system in terms of the functioning of its parts, assuming that these parts are performing discrete, isolatable sub-tasks. Decomposable systems exhibit division of labour—different components fulfil specified, nontransferable subtasks. Much of the original appeal of the mechanistic account came hand-in-hand with the componential account. The componential model inspired early scientists like Boyle, Descartes and Galileo (Dijskerhuis, 1950; Westfall, 1978). Nature was conceived as a system governed by a fine set of discrete entities that—by pure interaction—produced high-level phenomena. The distinction between primary and secondary qualities was crucial in this context. Primary qualities—size, solidity and extension—are properties that are intrinsic to the object in the sense

that they can be instantiated even if no other object is available in the world. Secondary qualities –smell, taste, distance—are, on the contrary, fundamentally dispositional. Mechanical philosophers gave explanatory prominence to primary qualities: they were supposed to be the ultimate building blocks that physical reality was made of. Secondary qualities, on the other hand, were considered explanatorily irrelevant: the product of the interaction of microscopic elements rather than the true cause of observable phenomena. The componential model conceived nature as a machine—as an automaton (Burtt, 1954 [2003]), and different technological inventions—the mechanical clock or the microscope—provided strong legitimacy to the mechanical project. The clock was able to perform complex, intelligent behaviour by performing blind, dumb movements (Mayr, 1986). The microscope in turn provided evidence of discrete microscopic organization at a level not visible by the human eye (Wilson, 1995).

Examples of decomposable models abound in science. In linguistics, for instance, semantics has been traditionally understood in compositional terms; the meaning of sentences is the result of the discrete meanings of their constituents. In psychology, Jerry Fodor argued that the mind could be decomposed in different, autonomous, encapsulate modules, each one processing a certain type of stimuli and delivering a certain type of output (Fodor, 1983). In the social sciences, perfectly decomposable explanations are rare, and most mechanistic explanations are of a hybrid nature, involving systems that are only partially decomposable.

I will briefly introduce an example of a hybrid mechanism in the social sciences. Suppose that we want to explain the emergence of leftist attitudes amongst intellectuals. Why do intellectuals tend to oppose the market? Philosopher Robert Nozick proposed the following explanation (Nozick, 1997). Intellectuals are highly educated individuals. They have not only spent many years within educational institutions, but they were also overrepresented amongst the best students of their cohort. Education implies prolonged contact with a system of distribution of rewards—the school—diametrically opposed to the market distributive principles. While school rewards are awarded according to a meritocratic principle, the market rewards are allocated on the basis of

marginal contribution. Intellectuals rapidly discovered that the selection operated by the market is not necessarily sensitive to scholastic performance: individuals that were very low on the scholastic ladder might be highly rewarded by the market. They regard this misalignment between schools and markets as unfair. The market is unfair simply because it only takes the endogenous preferences of economic agents into account to determining rewards, without using scholastic track records as a more objective and precise guide to allocate rewards.

Nozick's example is particularly interesting because it illustrates a system that is partially—but not fully—decomposable. Note that the educational system is assigned specific sub-tasks within this explanation. The educational system is responsible for stratifying individuals according to objective meritocratic principles. If the educational system does not function properly—failing to be meritocratic, for instance—the high level behaviour of the system (to wit, the tendency of intellectuals to regard the market as unfair) might not be produced. So schools are supposed to produce isolatable effects within the system. Although the system is still far from being completely decomposable, there are distinctive components that fulfil precise, discrete tasks (Ríos and Magni-Berton, 2003).

Furthermore there is a discrete psychological mechanism that can be separately tested—to wit the fact that school fosters expectations that the market betrays. The introduction of this psychological mechanism explains why intellectuals are unhappy with their incomes. Note that this psychological mechanism might be put to test in specific contexts, keeping the other factors involved relatively constants. One way to do this would be to gather data on the behaviour of intellectuals belonging to low-caste systems in India. The theory predict that low caste intellectuals would not regard the market as unfair, precisely because they did not have expectations of upward social mobility—an hypothesis suggested by Nozick's himself (Nozick, 1997).

Nozick's hypothesis is also interesting to illustrate the role of mechanisms within the economy of mechanistic explanations. There are in fact different functions that mechanisms perform and they should be carefully distinguished. To begin with, mechanisms are powerful tools to determine the direction of the causal relationship. In this way, to

postulate a mechanism contributes to the clarification of the semantic connection between the variables in play. The mechanism explains why a critical look at the market increases, not decreases, with education. Anti-market attitudes increase with education because the latter implies a prolonged and critical exposure to a distributive justice criterion that is diametrically opposed to those that rule the market. This interpretative function of mechanisms is fundamental for determining the direction of the causal relationship between the two variables involved in the explanation—attitudes vis-à-vis the market and educational achievement. It is the educational achievement—together with the luck of success in the market—that explains intellectuals' opposition to the market.

Mechanisms are also crucial for increasing the transparency of the system (Boudon, 1984; Cherkaoui, 2005; Elster, 2007). Providing a mechanism amounts to opening the 'black box' and presenting a detailed reconstruction of the process from 'inputs' to 'outputs'. The full explanation of the leftist ideology of intellectuals is achieved by breaking down the global relationship that exists between the variables into several sub-mechanisms. The final result is greater theoretical transparency and a maximum reduction of 'black boxes'. Discovering the underlying mechanism that rules a social event implies closing the enquiry. One can still wonder about the explanation of the mechanism itself, but not about the issue that has been the starting point of the investigation. Coming back to Nozick's example: it is certainly possible to ask why winners in the school system accept readily the meritocratic criterion. However, the obvious aim of such a question is the analysis of the mechanism itself, not the explanation of the emergence of anti-market attitudes amongst intellectuals. Depending on how deep we want to go, we might or we might not want to explore further inside the mechanisms themselves.

4. NON-DECOMPOSABLE SYSTEMS

In the previous section we briefly analysed the componential (and hybrid) versions of the mechanistic model. In this section I would like to turn to the alternative account—call it the non-decomposable model. In these cases, the high-level behaviour of the system is the result of the interaction of its components, but these components are not responsible

for performing discrete, nontransferable sub-tasks. The limiting case is that all these components are exchangeable: in a way all of them are doing the same thing. Their behaviour is homogeneous, and what is crucial is the relational interaction with the other components rather than their own intrinsic structure. Note that the behaviour of the system as a whole depends on the behaviour of its parts, but the sub-task performed by the sub-components is almost alike; there is no specialized, dedicated sub-components fulfilling nontransferable tasks. Systems like these are highly homogeneous, and they do not display modularity. This does not mean, of course, that the behaviour of the system does not depend on the behaviour of its parts: the high-level patterns of the system are still the result of the behaviour of their components.

A good example is the model proposed by Thomas Schelling to account for segregation (Schelling, 1969). Schelling placed pennies and dimes on a chessboard and moved them around according to various rules. He interpreted the board as a city, and individuals of two different groups (black and whites, boys and girls, etc.) with dimes and pennies. He distributed these pennies and dimes randomly across the board. Schelling later assumed that the pennies were "unhappy" if they had more than four dimes as neighbours, and the other way round, dimes were "unhappy" if they were surrounded by more than four pennies. Pennies having more than four dimes as neighbours, moved to a new location where there were fewer dimes, and the same occurred for dimes. The upshot of the process is the generation of segregation: dimes and pennies being concentrated in different parts of the chessboard. Note that the properties of the whole system are dependent on the behaviour of the components, but these components are homogeneously performing the same task—avoiding having more than four neighbours differing from themselves. Obviously there are different patterns that can be obtained in a purely mechanical way; each pattern will depend on the original distribution, the actual behaviour of the agents, the rules of the game and obviously the structure of the space. In this case, we assumed that the rules of the game were (1) if you are a penny, avoid having more than 4 dimes as neighbours; and (2) if you are a dime, avoid having more than 4 pennies surround you. There are different configurations that could emerge from this model, but all the

patterns will tend to have a segregated structure, with pennies clustering in one sector of the board, and dimes clustering in another sector.

Another possible example of a non-decompositionability is the construction of termite towers. Resnick suggests an extraordinary hypothesis to account for the building of the termite tower in a simple and elegant way (Resnick, 1997). It is in fact a mechanistic explanation based on massive decentralized processes: each termite performs an extremely simple action, which when combined with the identical actions of its peers, automatically produces a specific global result. Resnick's explanation is as follows: let us assume that we build a computer program to simulate the behaviour of a group of termites. First we create 10 termites. We then spread 100 grains of sand on the screen. We now want the termites to pick up the grains of sand and pile them up somewhere to create the termite tower. To do this, we implement two rules for the termites: (1) if you are a termite and you have a grain of sand on you, place it on another grain of sand; (2) if you are a termite and do not have a grain of sand, find one.

Resnick argues that these two rules generate the termite tower. The explanation for this surprising phenomenon is very simple. The first termite gets going, looks for a grain of sand—grain of sand number 100 for example—and places it on the grain immediately next to it—let us say number 99. This simple operation renders a drastic ecological change in the whole system: before the termite made this move, there were 100 possible places in which to build the termite tower, now there are only 99. Whatever the next termite does, it will have to place its grain of sand on top of another and the number of potential places will still—in the worst of cases—remain at 99, because one of the rules would otherwise have been broken. This means that even if the next termite separates the two grains of sand that the previous one piled up, it will have to place its own grain of sand on another grain of sand. We will never return to the situation where there were 100 places to build the termite tower.

The most important thing to note about non-decomposable systems is that they do not rely much on functional analysis. This is a striking difference with componential models. In componential explanations there is a lot of work that needs to be done to identify what each subcomponent of the system is doing; they are discrete, modular systems

having complex, dedicated sub-parts charged to perform specific tasks. Nothing like this happens in the case of non-decomposable models. There is not much to investigate in terms of the functions of each sub-component. The important issue in these systems is the interaction of the sub-components. The main focus is on understanding the logic of connectivity between the sub-components, rather than in describing the functional goals that each component of the system is performing.

The two models that we have analysed here are laboratory models. Obviously there are other non-decomposable models having a more practical bite. Connectionist models in psychology are also non-decomposable and they have been of crucial importance in the development of alternative frameworks in modern cognitive sciences, generating a wealth of hypotheses at variance with the classical symbolic approach (Bechtel & Abrahamsen, 1991; Betchel & Richardson, 1993). In economics, the neoclassical model of price fixation is close to being non-decomposable: the sub-components of the systems are essentially performing the same sub-tasks. Prices are fixed by the behaviour of each one of the economic agents; all of them are partially responsible for price fixation, but none of them is assigned a discrete, identifiable role within the economy. These systems cannot be broken down into parts having specific roles or tasks to fulfil. Each agent or component of the system is essentially like any other; they are all essentially doing the same thing. Obviously the more sub-systems you introduce to your model—think about institutions like central banks, for instance—the more modular a system will become, and consequently, the more it will depart from non-decomposability.

5. CODA

I have reviewed different issues concerning mechanistic explanations. I have shown why mechanistic explanations have not been discussed much in the traditional philosophy of science, despite their extraordinary importance in scientific practice itself. I have also distinguished different types of mechanistic explanations, depending on the degree compositionality exhibited by a system. I have analysed compositionality in terms of whether there are distinctive functions performed by

discrete, dedicated sub-components of the systems. The sub-components of non-decomposable systems exhibit low modularity and high interaction between the sub-components of the system. I have introduced different examples to try to illustrate these aspects of the mechanistic framework. Given the importance of the mechanistic model in modern science, and the epistemic richness of mechanisms-based explanations, the attention they have received in mainstream social sciences is not surprising. The development of the mechanistic model in the social sciences will greatly improve our understanding of the social world, leading us toward better and more accurate policy intervention.

REFERENCES

Akerlof, G.A. (1970) "The market for 'lemons': Quality uncertainty and the market mechanism", *Quarterly Journal of Economics*, 84(3): 488–500.

Bechtel, W. (2006) *Discovering Cell Mechanisms: The Creation of Modern Cell Biology.* Cambridge, MA: Cambridge University Press.

Bechtel, W., & Abrahamsen, A. (1991) *Connectionism and the Mind: An Introduction to Parallel Processing in Networks.* Oxford: Blackwell Publisher.

Betchel, W., & Richardson, R. (1993) *Discovering Complexity: Decomposition and Localization as Strategies in Scientific Research.* Cambridge, M.A.: MIT Press.

Bliss, M. (1984) *The Discovery of Insulin.* Chicago: The Chicago University Press.

Boudon, R. (1984) *La place du désordre.* Paris: Presses Universitaires de France.

Bunge, M. (1964) "Phenomenological theories", in M. Bunge, *The Critical Approach to Science and Philosophy.* London: Free Press.

Burtt, E. A. (1954 [2003]) *The Metaphysical Foundations of Modern Science.* Dover: Dover Publications.

Cherkaoui, M. (2005): *Invisible Codes: Essays on Generative Mechanisms.* Oxford: Bardwell Press.

Craver, C. (2007) *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience.* Oxford: Clarendon Press.

Dijskerhuis, E. (1950) *The Mechanization of the World Picture.* Princeton: Princeton University Press.

Elster, J. (2007) *Explaining Social Behavior: More Nuts and Bolts for the Social Sciences.* Cambridge: Cambridge University Press.

Fodor, J. (1983) *The Modularity of Mind.* Cambridge: MIT Press.

Giere, R. (1988) *Explaining Science: A Cognitive Approach.* Chicago: University of Chicago Press.

Hempel, C. (1965) *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science.* New York: Free Press.

Kuchle, G. and Ríos, D. (2012): "A game-theoretic analysis of the Baldwin Effect". *Erkenntnis,* 77(1): 31–49.

Mayr, O. (1986) *Authority, Liberty and Automatic Machinery in Early Modern Europe.* Baltimore, Maryland: Johns Hopkins University Press.
Nagel, E. (1961) *The Structure of Science.* New York: Harcourt, Brace and World.
Nozick, R. (1997) "Why do academics oppose capitalism?", in R. Nozick, *Socratic Puzzles.* Cambridge, MA: Harvard University Press.
Psillos, S. (2002) *Causation and Explanation.* Chesham: Acumen.
Resnick, M. (1997) *Turtles, Termites and Traffic Jams: Exploration in Massively Parallel Microworlds.* Boston, MA: MIT Press.
Ríos, D. (2004): "Mechanistic explanations in the social sciences", *Current Sociology* 52: 75–89.
Ríos, D. and Magni-Berton, R. (2003): *La misère des intellectuels*: *Pourquoi s'opposent-ils au capitalisme?* Paris: L'Harmattan.
Schelling, T. (1969) Models of Segregation. *American Economic Review , 59 (2)*, 488–483.
Simon, H. (1969) *The Sciences of the Artificial.* Cambridge, M.A.: MIT Press.
van Fraassen, B. (1980) *The Scientific Image.* Oxford: Oxford University Press.
van Parijs, P. (1981) *Evolutionary Explanation in the Social Sciences.* New Jersey: London & Totowa.
Weber, M. ([1905] 2009) *Die protestantische Ethik und der Geist des Kapitalismus.* Berlin: Anaconda.
Westfall, R. S. (1978) *The Construction of Modern Science: Mechanisms and Mechanics.* Cambridge: Cambridge University Press.
Wilson, C. (1995) *The Invisible World: Early Modern Philosophy and the Invention of the Microscope.* Princeton, New Jersey: Princeton University Press.
Wimsatt, W. (2007) *Re-Engineering Philosophy for Limited Beings: Piecewise Approximations to Reality.* Cambridge: Harvard University Press.

9

THE PLURALITY OF EXPLANATORY GAMES

C. Mantzavinos
University of Athens

1. INTRODUCTION

Three models are currently dominating the discussions in the theory of scientific explanation. First, the causal/mechanistic model which claims that an explanation consists in the identification of mechanisms understood as entities and activities organized such that they are productive of regular changes, from start to termination conditions (Salmon, 1984, Machamer *et al.* 2000). This is now the dominant approach in the social sciences (Hedström and Swedberg 1996, Elster 2007, Demeuleneare 2011). Mohamed Cherkaoui has been one of its protagonists, his position having been best articulated in his *Invisible Codes* (2005). Second, the unification approach, which claims that explanations are deductive arguments that provide understanding by fitting particular facts and events within a general theoretical framework (Friedman, 1974, Kitcher, 1981 and 1989, Bartelborth 2002). Finally, the manipulationist approach, which claims that explanatory activity consists mainly in answering the "what-if-things-had-been-different-question", i.e. it views an explanation primarily as enabling us to see what sort of difference it would have made for the explanandum if

the factors cited in the explanans had been different in various possible ways (Woodward 2003).

I have shown elsewhere (Mantzavinos, 2014) that these three approaches need not exclude one other, but they are compatible with each other, in the sense that each approach applies to different domains of science. Showing this kind of compatibility paves the way for the defence of the position of explanatory pluralism. This is a pluralistic account that allows for different ideal types of explanation, i.e. different exemplary accounts of classification of explanatory activities.

One fruitful way to illuminate explanatory pluralism is to acknowledge that at every moment of time there is a stock of explanations available in a society, proposed by ordinary people "in the wild" or by specialists organized formally or semi-formally within specific organizational structures such as churches, universities, etc. This explanatory reservoir is distributed among diverse individuals and groups in society under conditions of a cognitive division of labour. The terms of provision, control, and dissemination of explanations in this collective explanatory enterprise are regulated by the different rules that the participants have come to adopt over time. These rules incorporate the normative standards that guide the processes of the discovery and justification of explanations as well as the modes of their communication, dissemination, and adoption. They constitute *the rules of the explanatory game* that the participants are playing.

For the purposes of my present discussion, it is sufficient to state that an abstract characterization of an explanatory game can be given by the following classification of rules: 1) Constitutive rules, including a) rules determining what counts as an explanandum, b) rules determining what must be taken as given and c) rules determining the metaphysical presuppositions; 2) rules of representation. 3) rules of inference; and 4) rules of scope (Mantzavinos, 2013). My aim in this paper is rather to argue for the plurality of explanatory games. At every moment of time there is a plurality of explanatory games in a society that take place in parallel. In other words, it is not only within science that different explanatory games are played, but also within society at large. Lay people use rules of representation, rules of inference and rules of scope in order to provide explanations about phenomena that interest them.

Children permanently ask why-questions to their parents and receive responses that are aimed to satisfy their curiosity. The vast majority of explanations offered in a society are common-sense explanations by lay people, following quite rudimentary rules in their explanatory activities. However, these explanations very often, though by no means always, tend to change, and under appropriate circumstances they become more and more convincing. In other words, the explanatory labour is not only distributed among the different participants in the diverse games at a moment of time; they also evolve over time. This evolution takes place as the sets of rules that guide the activities of the explainers change over time. Thus, the broad picture is one of both a horizontal division of explanatory labour at a moment of time and a vertical division of explanatory labour over time. The horizontal division can be best conceptualized as a plurality of explanatory games existing in parallel at a time, and the vertical division as the rule-change of these games over time. I will elaborate further on this in what follows.

2. THE HORIZONTAL DIMENSION

At every moment of time there is a continuum of explanatory knowledge available in a society. This knowledge is produced while diverse communities (broadly defined to include any group of people) proceed in explanatory activities. Thus, many explanatory games take place concurrently. This is trivial, if one considers that *one set of the constitutive rules* of any explanatory game determines *what counts as an explanandum*. Human beings wish to explain a huge number of phenomena in their environment, so it is only natural to postulate that there, accordingly, must be a huge number of games. What makes the claim of plurality more interesting, however, is that the constitutive rules are also comprised by *rules determining what must be taken as given* and by *rules determining the metaphysical presuppositions* of the game. So, we can have explanatory games which have the same explananda, but differ in the rules determining what must be taken as given and/or what the metaphysical presuppositions are. Here the plurality becomes more interesting, because we can identify different explanatory games that deal with the same explananda, but which differ from one another on

the basis of in what they assume as given and of their metaphysical assumptions.

Three important types here are mythical, religious and scientific explanatory games. There are cases in which these three types of explanatory games simply deal with distinct and different explananda and, thus, there is no overlap among them (consider the question of why angels need to have wings in order to fly). In many cases, however, these three types of games deal with the same explananda, but differ radically in what they take as given and in their metaphysical presuppositions. Consider the case of the emergence of the universe, cosmogony.

A prominent explanation that was produced in the framework of the Greek mythical explanatory game is best summarized in the following quotation from Hesiod's *Theogony* (114–138).[1]

> Tell me these things, Olympian Muses, tell
> From the beginning, which first came to be?
> Chaos was first of all, but next appeared
> Broad-bosomed Earth, sure standing-place for all
> The gods who live on snowy Olympus' peak
> And misty Tartarus, in a recess
> Of broad-pathed earth, and Love, most beautiful
> Of all the deathless gods. He makes men weak,
> He overpowers the clever mind, and tames
> The spirit in the breasts of men and gods.
> From Chaos came black Night and Erebos.
> And Night in turn gave birth to Day and Space
> Whom she conceived in love to Erebos.
> And Earth bore starry Heaven, first, to be
> An equal to herself, to cover her
> All over, and to be a resting-place,
> Always secure, for all the blessed gods.
> Then she brought forth long hills, the lovely homes
> Of goddesses, the Nymphs who live among
> The mountain clefts. Then, without pleasant love,
> She bore the barren sea with its swollen waves,
> Pontus. And then she lay with Heaven, and bore

Deep-whirling Oceanus and Koios; then
Kreius, Iapetos, Hyperion,
Theia, Rhea, Themis, Mnemosyne,
Lovely Tethys, and Phoebe, golden-crowned.
Last, after these, most terrible of sons,
The crooked-scheming Kronos came to birth
Who was his vigorous father's enemy.

A prominent explanation of cosmogony that was produced in the framework of the Christian explanatory game is mainly contained in the Book of Genesis, the first chapter of which I would like to quote at length:

> In the beginning when God created the heavens and the earth, the earth was a formless void and darkness covered the face of the deep, while a wind from God swept over the face of the waters. Then God said, "Let there be light"; and there was light. And God saw that the light was good; and God separated the light from the darkness. God called the light Day, and the darkness he called Night. And there was evening and there was morning, the first day. And God said, "Let there be a dome in the midst of the waters, and let it separate the waters from the waters." So God made the dome and separated the waters that were under the dome from the waters that were above the dome. And it was so. God called the dome Sky. And there was evening and there was morning, the second day. And God said, "Let the waters under the sky be gathered together into one place, and let the dry land appear." And it was so. God called the dry land Earth, and the waters that were gathered together he called Seas. And God saw that it was good. Then God said, "Let the earth put forth vegetation: plants yielding seed, and fruit trees of every kind on earth that bear fruit with the seed in it." And it was so. The earth brought forth vegetation: plants yielding seed of every kind, and trees of every kind bearing fruit with the seed in it. And God saw that it was good. And there was evening and there was morning, the third day. And God said,

> "Let there be lights in the dome of the sky to separate the day from the night; and let them be for signs and for seasons and for days and years, and let them be lights in the dome of the sky to give light upon the earth." And it was so. God made the two great lights—the greater light to rule the day and the lesser light to rule the night—and the stars. God set them in the dome of the sky to give light upon the earth, to rule over the day and over the night, and to separate the light from the darkness. And God saw that it was good. And there was evening and there was morning, the fourth day.

Contemporary cosmology, the branch of physics that deals with the question of the emergence of the universe, proposes a big bang explanation. According to the general theory of relativity, gravity is not a force, but the curvature of space time. The range of gravity is infinite since it is a property of space time itself and the evolution of the universe is ultimately dictated by gravity. A set of differential equations relates the dynamic quantities in the universe and turn them back in time. Simulations on computers relate initial conditions of the universe to different outcomes aiming at representations of the kind of universe we live in. Earlier astronomers, most prominently Edwin Hubble, observed that neighbouring galaxies of the Milky Way galaxy are receding away—the more distant they are, the faster were they found to be moving away. In other words, recessional velocity of a galaxy increases with its distance from the earth. Reversing this expansion scenario back in time, the inference is drawn that if galaxies are moving rapidly apart now, they must have been denser—with more matter and energy per unit volume—in the past. Going back in time, the whole universe should converge to a point of infinite density and extremely high temperatures, and this should be the starting point of big bang. The universe itself should have been an infinitely dense point which expanded to its present size. The age of the universe according to this Big Bang explanation is estimated to be 13,73 billion years.

Given that Hesiod was active between 750 and 650 BC and that his mythical explanation has been reproduced for centuries afterwards and that Genesis was composed in the late 7th or the 6th century BC

(according to the prevailing consensus of biblical scholars) the first two explanatory games have been unfolding in parallel for a long period of time. Today, nobody believes the mythical explanation of Hesiod to be true. Today, it is the religious explanatory game of the Genesis, along with the scientific explanatory game of cosmology, that are unfolding in parallel. (And there are other religious explanatory games that deal with the same explanandum, which I leave unmentioned here).

The biblical and the scientific explanatory game are structured around the same explanandum, but *differ radically in the rules determining what must be taken as given and the rules determining their metaphysical presuppositions*. In their everyday work, though the players of the respective games engage in their explanatory activities following the constitutive rules of their own game, they occasionally engage critically with the constitutive rules of the other games. This engagement is usually manifest in the rejection of what is taken as given by the players of the other game and a rejection of the metaphysical presuppositions that they employ. The ongoing debate on creationism is such an example.

3. THE VERTICAL DIMENSION

Explanatory games evolve over time. An important factor of their evolution is that even the most fundamental of the constitutive rules, the ones that determine what counts as an explanandum, evolve over time. The legitimate why-questions to ask in a specific domain change over time, something that should not be astonishing.[2]

The change in the constitutive rules is certainly important, but the *prevalent change of all other kinds of rules* is of equal importance, making the playing of an explanatory game an inherently dynamic enterprise. *The means of representing phenomena* evolve over time—in contemporary science, artefacts like graphs, computer monitor displays, etc. are constantly developed further as are abstract means of representation like mathematical models. The same is the case with *the rules of inference*: this is what the traditional philosophy of science has focussed on, debating the nature of these rules, i.e. whether they are lawful, and if so, what their lawfulness consists in, whether they function as mere inference tickets from data to data or whether one should interpret them

realistically, whether the different ontological structures require different kinds of law-like generalizations, etc. Finally, *the rules of scope* also change over time: new instructions about where and how to apply the explanatory rules to new phenomena are issued, debated and followed by the participating scientists.

The explanations offered by the participants of an explanatory game evolve over time as they follow the different kinds of rules available to them. Explanatory knowledge is just a part of general human knowledge that is a product of genetic and cultural evolution. The process of cultural evolution concerns the change and transmission of knowledge in time at a societal level, and it can be regarded as a process of collective learning (Mantzavinos, 2001, p. 73). Donald (1991), in his *Origins of the Modern Mind*, highlights the function of what he calls "External Symbolic Storage" for the transmission and accumulation of knowledge across generations. It was the simple habit of recording ideas, that is, "of externalizing the process of oral commentary of events" (p. 342) that constituted the critical innovation that has massively supported the evolution of theoretic culture. This occurred in Greece beginning around 700 BCE. What was truly innovative was that "for the first time in history complex ideas were placed in the public arena, in an external medium, where they could undergo refinement over the longer term, that is, well beyond the life-span of single individuals" (p. 344). These External Symbolic Storage Networks have decisively changed the character of knowledge evolution, since they provided the possibility of a constant interaction between the accumulated corpus of theoretical knowledge and the theoretical problems of the individuals in a society.

If one adopts an evolutionary perspective, then the thesis of a continuum between everyday, commonsensical knowledge and scientific knowledge, both products of the more general cultural evolutionary process, seems very plausible. The thesis of a continuum does not deny the difference in the quality of the scientific knowledge vis à vis other kinds of knowledge. But instead of focussing on clear-cut demarcation criteria of a *syntactic nature* (something that was fashionable for a long time in philosophy of science), it suggests that *it is the specific way that knowledge is imposed for criticism that one should rather study*. An institutional arrangement that enables criticism is the first prerequisite, of

course. With this I mean that the institutional structure, both in its formal and in its informal elements, must allow or even encourage the criticism of knowledge structures and practices. A polity that guarantees the freedom of expression and allows for resources to be devoted to the building of organisational structures like universities, where everything is allowed to be questioned, provides the formal institutional structure that makes advanced criticism possible. A widespread critical attitude towards given beliefs and practices provides the informal institutional structure that enables criticism. The latter is more important (and has in fact come first historically).[3] The specific mix of formal and informal institutions prevailing in a society regulates the behaviour of the individuals and provides the solution to what is (very inaccurately called) the problem of "power".[4]

In theocratic systems there is strong resistance to the criticism of religious explanations. Communist political systems have similarly regularly disallowed the development of explanatory games for the analysis of social phenomena other than the ones that the communist party has found acceptable. The range and amount of criticism thus decisively depends on the prevailing institutional framework—the freedom of thought and expression that enables scientific explanatory activities in modern universities and is guaranteed by liberal states around the world is only a contingent historical case.

4. EPILOGUE

Thus, explanatory games unfold through time in the context of specific institutional constraints, and in the broader context of an ongoing evolution of all human knowledge. The rules of representation, the rules of inference and the rules of scope followed by the participants of any specific game change over time. Since the ingredients of explanations come from the use of these different kinds of rules, their change over time is concomitant with a vertical division of explanatory labour. The broad picture of the division of explanatory labour consists, thus, both of a horizontal dimension comprising all explanatory games at a moment of time and of a vertical dimension encapsulating all changes of explanatory activities taking place within the diverse games over time.

Working out exactly how these rules change over time can not be taken up here. That clearly remains a desideratum.

NOTES

1. This the translation in the Penguin Classics edition (1976). For a useful discussion of cosmogonic myths see Burkert (1999).
2. Thomas Kuhn (1962/1970, p. 106ff.) has drawn attention to that and van Fraassen (1980, p. 111f) has raised the issue again in the theory of explanation: "Examples [Kuhn] gives of explanation requests which were considered legitimate in some periods and rejected in others cover a wide range of topics. They include the qualities of compounds in chemical theory (explained before Lavoisier's reform, and not considered something to be explained in the nineteenth century, but now again the subject of chemical explanation). Clerk Maxwell accepted as legitimate the request to explain electromagnetic phenomena within mechanics. As his theory became more successful and more widely accepted, scientists ceased to see the lack of this as a shortcoming. The same had happened with Newton's theory of gravitation which did not (in the opinion of Newton or his contemporaries) contain an explanation of gravitational phenomena, but only a description. In both cases there came a stage at which such problems were classed as intrinsically illegitimate, and regarded exactly as the request for an explanation of why a body retains its velocity in the absence of impressed forces. While all of this may be interpreted in various ways (such as through Kuhn's theory of paradigms) the important fact for the theory of explanation is that not everything in a theory's domain is a legitimate topic for why-questions; and that what is, is not determinable *a priori*.)".
3. See Popper (1963/1989, p. 50): "For the critical attitude is not so much opposed to the dogmatic attitude as super-imposed upon it: criticism must be directed against existing and influential beliefs in need of critical revision—in other words, dogmatic beliefs. A critical attitude needs for its raw material, as it were, theories of beliefs which are held more or less dogmatically. Thus science must begin with myths, and with the criticism of myths; neither with the collection of observations, nor with the invention of experiments, but with the critical discussion of myths, and of magical techniques and practices. The scientific tradition is distinguished from the pre-scientific tradition in having two layers. Like the latter, it passes on its theories; but it also passes on a critical attitude towards them. The theories are passed on, not as dogmas, but rather with the challenge to discuss them and improve upon them. This tradition is Hellenic: it may be traced back to Thales, founder of the first school (I do not mean 'of the first philosophical school', but simply 'of the first school') which was not mainly concerned with the preservation of a dogma."
4. I have developed my own theory of institutions which gives what I think a more accurate account of how "power"—indeed a very vague term and thus, a useless one—is regulated in a society. See Mantzavinos (2001).

REFERENCES

Bartelborth, Thomas (2002) "Explanatory unification", *Synthèse*, 130: 91–107.

Bible, Old Testament.

Burkert, Walter (1999) "The logic of cosmogony", in Richard Buxton (ed.), *From Myth to Reason? Studies in the Development of Greek Thought*. Oxford: Oxford University Press, pp. 87–106.

Cherkaoui, Mohamed (2005) *Invisible Codes. Essays in Generative Mechanisms*. Oxford: The Bardwell Press.

Demeulenaere, Pierre (ed.) (2011) *Analytical Sociology and Social Mechanisms*. Cambridge: Cambridge University Press.

Donald, Merlin (1991) *The Origins of the Modern Mind. Three Stages in the Evolution of Culture and Cognition*. Cambridge/Mass.: Harvard University Press.

Elster, Jon (2007) *Explaining Social Behavior. More Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

Friedman, Michael (1974) "Explanation and scientific understanding," *Journal of Philosophy*, 71: 5–19.

Hedström, Peter, and Richard Swedberg (eds) (1996) *Social Mechanisms. An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.

Hesiod (1976) *Theogony*. London: Penguin Classics.

Kitcher, Philip (1981) "Explanatory unification", *Philosophy of Science*, 48: 251–281.

——— (1989) "Explanatory unification and the causal structure of the world", in Philip Kitcher and Wesley Salmon (eds), *Scientific Explanation*, volume 13 of Minnesota Studies in the Philosophy of Science, pp. 410–505.

Kuhn, Thomas (1962/1970) *The Structure of Scientific Revolutions*, second enlarged edition. Chicago: Chicago University Press.

Machamer, Peter, Lindley Darden, and Carl Craver (2000) "Thinking about mechanisms", *Philosophy of Science*, 67: 1–25.

Mantzavinos, C. (2001) *Individuals, Institutions and Markets*. Cambridge: Cambridge University Press.

——— (ed.) (2009) *Philosophy of the Social Sciences. Philosophical Theory and Scientific Practice*. Cambridge: Cambridge University Press.

——— (2013) "Explanatory games", *Journal of Philosophy*, 110: 606–632.

——— (2014) "Explanation" in: *International Encyclopedia of the Social and Behavioral Sciences*, 2nd edition (forthcoming).

Popper, Karl R. (1963/1989) *Conjectures and Refutations. The Growth of Scientific Knowledge*, 5th revised edn. London: Routledge.

Salmon, Wesley (1984) *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.

van Fraassen, Bas C. (1980) *The Scientific Image*. Oxford: Clarendon Press.

Woodward, James (2003) *Making Things Happen*. Oxford: Oxford University Press.

10

REDUCTION THAT REALLY MATTERS

Can We 'Reduce' from Macro to Micro in the Social Sciences (And Should We)?

José A. Noguera
Grupo de Sociología Analítica y Diseño Institucional (GSADI),
Department of Sociology, Universitat Autònoma de Barcelona

1. METHODOLOGICAL INDIVIDUALISM, REDUCTION, AND ANALYTICAL SOCIOLOGY

Prejudices on what others think are usually widespread in society, and also among scholars in the social sciences. For instance, non-analytical sociologists tend to think that all analytical sociologists are methodological individualists and advocates of rational choice theory. They would be astonished to see how much criticism against those options is launched in analytical sociology conferences (as I was when I first came into the analytical sociology circles). It could in fact be claimed that the harder and deeper critiques to methodological individualism[1] have come from those who are supposedly their most dogmatic defenders in the eyes of the rest of the sociological community (this is hardly realized, for example, among Pierre Bourdieu's followers).

Surprising as that may be, the fact is that the inspiring 'fathers' of analytical sociology such as Raymond Boudon, Jon Elster, or James

S. Coleman have very clearly subscribed to one or another version of 'methodological individualism' (MI henceforth). A well-known and useful definition of MI is the one given by Elster: "explanations in the social sciences should refer only to individuals and their actions" (2007: 13); in an earlier work, he proposed a more detailed definition which reads like this:

> By methodological individualism I mean the doctrine that all social phenomena (their structure and their change) are in principle explicable only in terms of individuals—their properties, goals, and beliefs. This doctrine is not incompatible with any of the following true statements. (a) Individuals often have goals that involve the welfare of other individuals. (b) They often have beliefs about supra-individual entities that are not reducible to beliefs about individuals. "The capitalists fear the working class" cannot be reduced to the feelings of capitalists concerning individual workers. By contrast, "The capitalists' profit is threatened by the working class" can be reduced to a complex statement about the consequences of the actions taken by individual workers. (c) Many properties of individuals, such as "powerful," are irreducibly relational, so that accurate description of one individual may require reference to other individuals. (Elster, 1982: 453)

So defined, I always thought MI was a trivial truth, as Elster himself has often noted. Why has it been, then, the object of so much harsh criticism and disdain? One of the accusations most frequently thrown against MI is that it is a reductionist strategy. I think this is right (Elster does too: see 2007: 13, 257–259): reduction consists on explaining macro-level phenomena in terms of micro-level, more elemental phenomena, so to explain macro-social phenomena in terms of individual behaviour is a reductionist strategy. However, I confess I share the naive belief that scientists are in the business of reduction (among others) and that if science is worth the effort, then reduction is a good thing. I share Elster's statement that "reductionism is the engine of progress in science" (2007: 258).

It is true that the words 'reduction' and 'individualism' have been given so many different meanings and interpretations that disputes

around them are highly poisoned by deep misunderstandings. It is probably not very intelligent to still try to 'save' them in no-matter-how nuanced or deflationary ways. However, I do not have better terms to express an idea which, to my view, is sound and trivially true; so, since I am no friend of changing conceptual labels in science for marketing reasons, I will rather stick to those terms here.

That MI is a discredited label may be shown by the fact that even analytical sociologists have felt the need of differentiating themselves from its supposedly 'crude' version in Elster's work, by saying they are rather 'structural individualists'; however, one might have trouble to see what is the conceptual difference between Elster's MI and the definition of 'structural individualism' as formulated in *The Oxford Handbook of Analytical Sociology* (2009):

> structural individualism is a methodological doctrine according to which all social facts, their structure and change, are in principle explicable in terms of individuals, their properties, actions, and relations to one another (Hedström and Bearman 2009: 8).

To this point, the definition seems clearly similar to the one previously quoted by Elster; to say the least, there is no incompatibility whatsoever between both definitions. But the authors immediately add:

> It differs from traditional notions of methodological individualism (e.g. Elster 1982) by emphasizing the explanatory importance of relations and relational structures (Hedström and Bearman 2009: 8).

And, only a page later, they state:

> That relations are important for explaining outcomes does not mean that they are independent of individuals and their actions, however. As emphasized above, in principle all relational structures are explainable as intended or unintended outcomes of individuals' action.

Again, it is clear that no logical incompatibility between this 'emphasis' on relations and Elster's definition of MI exists: point *(c)* of the above quote from 1982 states without doubt that in Elster's view relations among individuals and relational properties (and, therefore, 'relational structures') are legitimate candidates to be included in the *explanans* by social scientists. Perhaps realizing this, Hedström and Bearman present the 'difference' just as a stronger 'emphasis' on relations. It may be surprising, though, that a question of 'emphasis', which is thus implicitly denied to have a conceptual or logical status, justifies the adoption of a new conceptual label. However, as suggested above, I realize that it may be a smart move of 'frame' in the context of the present discredit and widespread prejudice against MI.

'Reductionism' has also been largely criticized and is often used as a pejorative label which dispenses from further arguments against an author or position.[2] Within the analytical circles, this feeling is also growing. Take for example the recent and insightful contributions by Daniel Little (2012) and Gianluca Manzo (2012) on analytical sociology: Little defends a so-called 'methodological localism' as an alternative to classical MI and to 'structural individualism', and one of the motivations to do that is to avoid 'reductionism' (Little, 2012: 149). Manzo brilliantly argues that 'methodological localism' is similar in any relevant sense to 'structural individualism', and he aims to refute the charge of 'reductionism' on analytical sociology. Both, therefore, seem to depart from the assumption that 'reductionism', whatever it is, is a bad thing.[3]

So did Mohamed Cherkaoui in Chapter 5 of his *Invisible Codes* (2005), entitled 'The Multiple Levels of Reality', and subtitled 'Is the macro always based on the micro?'. There, the following statement can be found:

> Social actionists rediscovered the idea dear to Weber and Popper that all social phenomena (macroscopic level), particularly the structure and function of institutions, should be explained and understood as the result of individual actions, not in terms of universals or collective properties. But in affirming this epistemological principle, actionist theorists did not reduce social phenomena to psychological ones.

Cherkaoui was depicting here a very usual view among sociologists in the analytical tradition, and even beyond: a sort of broad anti-reductionist individualism, allegedly not to be mixed up with strict MI.

In what follows, I will schematically argue two points: first, it is not especially problematic in sociology to endorse 'reductionism' in a very simple (and maybe trivial) but important sense: ontological reduction. Second, contrary to the prevalent view, ontological and methodological individualism may have little sense as independent positions.

2. WHAT DOES 'REDUCTION' MEAN?

Reduction as a Scientific Virtue

'Reduction' is, as suggested above, a term that has been given many different meanings. Of course, as remarked by Elster, different versions of reductionism (or any other research strategy) may be 'bad' and 'good', and there is no point in stating that 'bad reductionism' is a bad thing (Elster 2007: 259). But most critics seem to deny the very possibility of any 'good' reductionism in social science. Whether they are 'individualist' (under any of the existing versions) or not, most social scientists seem to agree that reduction is an unfeasible and/or dangerous aim.

However, one may argue that reductionism should be a good thing for science. The reductionist stance leads to study the world as a set of assembled parts or elements that may be disassembled (in a material sense or using mental abstraction) and studied separately. A reductionist explanation of a given phenomenon tries to clearly identify its constitutive parts and to analyse how the properties of these parts and the interaction between them have produced the phenomenon. Reductionism, so understood, is one of the fundamental pillars of modern science, as Edward O. Wilson nicely reminds us:

> The cutting edge of science is reductionism, the breaking apart of nature into its natural constituents. The very word, it is true, has a sterile and invasive ring, like scalpel or catheter. Critics of science sometimes portray reductionism as an obsessional disorder, declining toward a terminal stage one writer recently dubbed "reductive megalomania." That characterization is an

> actionable misdiagnosis. Practicing scientists, whose business is to make verifiable discoveries, view reductionism in an entirely different way: It is the search strategy employed to find points of entry into otherwise impenetrably complex systems. Complexity is what interests scientists in the end, not simplicity. Reductionism is the way to understand it. The love of complexity without reductionism makes art; the love of complexity with reductionism makes science' (Wilson 1998: 58–59).

Notice that this is not equating reduction and explanation. Reduction shows that some entities are composed of other more elementary entities, and that this fact is relevant for explaining how they are produced and why they exist. Here "is relevant for" is important: the reduction is not the explanation by itself, but it helps to explain. You have not explained a mob when you show that its reduction base are individual actions or mental states; you did it when you showed which exact composition of actions and mental states, causally connected in an specific way, gave birth to the mob behaviour, as for example in Mark Granovetter's (1978) threshold models of collective behaviour. If you know the reduction base, you still have to work on how to explain the phenomenon departing from that base; but if you do not know the reduction base, your explanation will be incomplete or false.

Reductionism is a scientific virtue because it *improves intelligibility* in scientific explanations. To reject reduction as an aim in a scientific discipline is to accept that there are ontological gaps in reality, or that some phenomena may be self-caused or constituted out of nothing, as if they were separated from the network of causal and/or constitutive relationships that connects all phenomena. To use Steven Weinberg's expression (1992: 56), when a particular 'reduction' is achieved in science, we are a little bit closer to the point where all our explanations will ideally converge. Our explanations have 'arrows' that point to places where we can keep on asking 'why'; reduction helps to successfully go down this chain of 'whys', by making disappear the gaps in it.

Is this also true for the social sciences? At first sight, one would say that if they made substantial progress along reductionist lines, this could only be considered as a progress in their scientific maturity. However,

many social scientists since Émile Durkheim have feared that any reductionist achievement would be a step towards the elimination of their disciplines. But is *reduction* necessarily accompanied by *elimination*? It does not seem so.[4] We may just observe that if many sociologists' fear or reductionism was justified, lots of chemists or biologists would have lost their jobs years ago to give way to physicists. Although reduction ideally aims to the convergence of all explanatory arrows into a common source, that ideal does not mean that all scientific disciplines except one must disappear, because, among other reasons, many of these arrows may depend on knowledge provided by different disciplines. The most reasonable and reduction-friendly model for the integration of scientific disciplines is not elimination but *consilience* (Wilson, 1998): the interconnection between events and theories based on facts studied by different disciplines, so that a common explanatory ground is generated.

Concepts of 'Reduction'

However, the answer to the question on elimination depends on the meaning we ascribe to the concept of 'reduction'. It is true that reduction in the sense described so forth (the Elster–Wilson–Weinberg sense) does not entail elimination. But, among many others, Lars Udehn claims in an influential book that scientific reduction "means elimination without remainder" (2001: 324–325). The point is he is using a different concept of 'reduction'. So what may 'reduction' mean? Which version of 'reductionism' is most suitable for the social sciences? Which are the *reducibility criteria* that a reductionist social science should satisfy? It is crucial to answer this question, since, as noticed by Jaegwon Kim (1999: 13), most of the objections against reductionism assume empty or inadequate models of what reduction entails in real scientific practice. Here I will broadly rely on John R. Searle (1992: 122–126) in order to distinguish between three types of reduction: *ontological*, *theoretical*, and *definitional*.[5]

a) *Ontological reduction*. What I will call, for simplicity, ontological reduction is a mix of two claims. First, an ontological one: objects of certain types are nothing but configurations of objects of other types. Second, a causal or methodological one: the causal powers of a given

entity are explainable in terms of the causal powers of other entities which conform it. I realize that it is highly polemical to merge both kind of thesis into a single concept of 'reduction', since it is very usual to agree with ontological reduction but not with methodological one (or the reverse). However, I have problems to accept this separation between ontology and methodology, so I will deal with this specific issue in the next section. Here it will suffice to say that, according to Searle, "the most important form of reduction is ontological reduction", since "it is the form of reduction to which the other forms point to" (1992: 123). In a similar way, Ulises C. Moulines has remarked that "the concepts of reduction and emergence clearly appear to have primarily an ontological meaning" (2006: 314). I think this is also the case for the social sciences.

The fact that ontological reduction is the most interesting form of reduction from a *scientific* point of view is fully compatible with the fact that it may be *philosophically* trivial, because it does not pose important philosophical problems. Instead, theoretical and definitional forms of reduction are the ones that usually produce philosophical problems.

b) Theoretical reduction. We have a theoretical reduction when all laws and propositions in a theory or discipline are *derivable* from the laws and propositions of other more fundamental theory or discipline; in such a case scientists often say that the first theory or discipline has been reduced to the second, or simply eliminated. However, this is a very unusual case in science (Searle, 1992: 123–124). In a classical article, Philip Anderson (1972), showed the difficulty of generalizing this requirement to all scientific theories, because the aspiration to derive all laws and scientific facts in the world from a simple set of 'ultimate' principles or equations generates enormous *computability* problems given a certain level of complexity. The famous Laplace's demon, hypothetically able to undertake such a project, is nothing more than a mental experiment.

One way of avoiding usual misunderstandings is therefore to distinguish the requirement of theoretical reduction (which, to use Anderson's term, could be regarded as a sort of logical *constructionism* rather than reductionism) from ontological reduction. This distinction helps to see that systematic non-derivability of scientific propositions at the macro level from those at the micro level in the social sciences, as

well as in many others, does not entail the impossibility to ontologically reduce the latter to the former; and this second reduction is the one that really matters in science.

c) Definitional (or linguistic) reduction. While ontological reduction has to do with the nature of the phenomena under study (and therefore with their proper description and explanation), and theoretical reduction has to do with the logical possibility of deriving some propositions from others, definitional reduction is about the possibility of *translation* between two different 'languages' or 'vocabularies'. Definitional reduction, according to Searle (1992: 124), is unfashionable, since it only cares about the relations between words or sentences, and how those referring to some kind of entities may (or not) be translated 'without information loss' to others referring to different kinds of entities. There are many cases where this is possible. Trivially, when definitional reduction is possible, then ontological reduction should also be possible. But the reverse is not true: the possibility of ontological reduction (the one which really matters) does not necessarily imply the possibility of definitional reduction (nor, even less, the need for it); only by going back to an atomistic conception of language such as the one shared by logical positivists could we support that implication. The impossibility of definitional reduction in a given case cannot then be a valid argument for the impossibility of ontological reduction.

It is however very usual in the social sciences to share an anti-reductionist stance based on the claim that in order to adequately explain their objects social scientists need concepts which refer to macro-social phenomena and which would be impossible, hard, or inconvenient to depict in terms of their microfoundations (such as systemic properties, institutional structures, cultural patterns, and the like).[6] To my view, this approach fails to make the distinction between definitional and ontological reduction, and can only appear to be 'anti-reductionist' because it focuses only on the definitional version of reduction. But Max Weber already noticed that the sociological interpretation of actions "is frequently forced to work with such concepts [that is, which 'supraindividual' concepts such as 'state', 'firm', 'organization', 'bureaucracy', and the like, JAN] (…) with the aim of achieving an intelligible *terminology*"

(Weber, 1922: 12); besides, Weber argued, though it would be possible to eliminate these concepts from usual language by introducing long-winded and wordy descriptions of their microfoundations, we still would face the fact that individual actors think of them as real entities, and that thought may have causal effects on their conduct.

This is why Elster's definition of MI soundly recalls that trying to reduce individuals' beliefs in supraindividual entities is a category mistake. As any other ontological reduction, MI operates in *extensional* contexts, not in *intensional* ones (Elster, 1985: 6).[7] The contrary would force us to treat as *real* any supraindividual or collective entity social actors may *believe* in, simply because such belief may have a causal effect on their behaviour. As wisely observed by Ronald P. Dore (1961: 76), claiming that social scientists should take their concepts directly from their data (such as individuals' beliefs) would be as absurd as forcing the carpenter to use only wooden saws. The truth conditions of the social scientists' explanatory propositions which include agents' beliefs are different from the truth conditions of those beliefs. 'John believes that a collective conscience led him to do action A' is true or false independently of the truth or falseness of 'A collective conscience led John to do action A'.

It seems to me, then, that social scientists could agree without important problems that ontological reduction is a good scientific practice; that it is the form of reduction that really matters in science, and that MI, understood in this sense, is the most obvious form that a reductionist strategy may take in the social sciences.

3. ONTOLOGY VS. METHODOLOGY: A DUBIOUS DISCONNECTION

A seemingly conclusion of the preceding section is that the real discussion about MI and reductionism is an ontological one. Why, then, to speak about 'methodological' individualism? Although the question is legitimate, the issue is purely terminological, and does not affect the core discussion of whether the research strategy depicted in Elster's definition is acceptable or not, irrespective of how we call it. We could change

the label to 'ontological individualism' and this, I think, would change little about the validity of the claims made.

But my aim in this section is to go beyond terminological discussions. In the previous section I defined ontological reduction as a merge of causal-methodological and ontological claims. However, a very widespread view tends to sharply separate both levels. I have always been puzzled to see how many philosophers and social scientists since Durkheim explicitly or implicitly claim to be ontological individualists but not methodological ones (see as a good example, Little, 2012). The reverse is also true: how come that so many social scientists have been methodological individualists (or reductionists) but not ontological ones? (think for example on many neoclassical economists who implicitly use MI while avoiding any ontological commitment, in a typical 'Friedmanite' fashion).

Here is the question I want to consider: once a stand has been taken for methodological (ontological) individualism, should not one go all the way through and embrace also ontological (methodological) individualism? I tend to think the answer is 'yes' in both directions. The reason is that the price to pay otherwise is unrealism, and that is a high price indeed for those social scientists who at the same time criticize rational choice theory for that reason.

Suppose that you claim that ontological individualism is sound ('of course social reality is made up from individual actions'), but that the explanatory strategy of methodological individualism is not. Suppose that, at the same time, you favour the typical analytical sociology objections against rational choice theory, and accuse that theory of lack of realism and indifference to the real mechanisms in place behind individuals' behaviour. There is some inconsistency in giving priority to methodology in order to criticize MI, but to ontology in order to criticize rational choice theory. You simply cannot have both at the same time, at least without making your point further. If a social scientist thinks that for methodological reasons it is sound to build explanations that ignore his ontological beliefs about how society is made up of individual actions, then he (more than other social scientists) should be sensitive to Milton Friedman's (1966) defence of unrealistic assumptions in economic theory. He may not like the particular unrealistic

assumptions Friedman (or any other individualist) make, but he simply cannot criticize those assumptions for their being unrealistic, since he is an ontological individualist but not a methodological one, so he implicitly admits that the ontological components of social reality (individuals) are of little explanatory interest for him.

Suppose now that you are, like Friedman, a methodological individualist, but not an ontological one ('modelling economic phenomena as if they were caused by individual decisions is sound even when I am making no ontological claim that individuals really decide anything, or even exist'). Then realism is also sacrificed, and the door is open for any methodological fiction that approximately predicts some pattern or helps our models to 'fit' the data.

In the first case (OI but not MI) one would be saying: 'I know that social reality is like this but I prefer to represent it like that'. In the second case (MI but not OI), one would be saying: 'I prefer to represent social reality like this, whether it is like this or not (which I do not know or care)'. I may be wrong, but I have always had a hard time trying to understand why someone committed to scientific realism might ever be sympathetic to any of those positions. To say it differently, it is hard to consider as totally independent things the methodological units of our scientific analysis and the really existing entities in the world. Would it not be more sensible to assume that, if one supports scientific realism, his methodological and ontological commitments should go hand-in-hand, also in the case of individualism?

4. CONCLUDING REMARKS: WHO'S AFRAID OF REDUCTIONISM?

Let us go back to Cherkaoui's remark quoted in section one: classical action theorists like Weber were individualists (they thought "that all social phenomena (…) should be explained and understood as the result of individual actions"), but they "did not reduce social phenomena to psychological ones". How should we understand this sentence? I cannot dare to know what Cherkaoui exactly meant by 'reduce', but I would bet it was not ontological reduction as it has been defined above, since that would be inconsistent with the mandate of explaining and

understanding social reality in terms of individual actions. So what does 'reduction' mean here? My guess is that it may mean theoretical reduction (sociology was not to be eliminated by psychology according to those classic theorists) or definitional reduction (the vocabulary of social-supraindividual entities would not be replaced by a collection of wordy descriptions in psychological terms). If, on the contrary, Cherkaoui was ascribing methodological individualism to Weber, but not ontological one, as well as supporting that option himself, then they would face the problem I dealt with in section three about the implausibility of a sharp separation between the ontological and the methodological levels, at least for scientific realists like them.

Be that as it may, and leaving aside many other relevant aspects of the debate on reductionism and MI, contemporary social science seems in his way to achieve some success on the lines of reduction to lower-level phenomena. This growing success could have the side-effect of reducing epistemological and methodological fragmentation, and enhancing the production of social-scientific knowledge in a controlled and accumulative fashion, something that has been more the exception than the rule in sociology. I have tried to present some thoughts on why sociologists (and, in general, social scientists) should be more sympathetic to reductionism and MI in the defined sense than they currently are. To be sure, it will be difficult to make disappear all misunderstandings, resistance and prejudices against this research strategy. But I do think it is still worth trying.

NOTES

The Spanish National Plan for R&D has supported this work through the CONSOLIDER-INGENIO 2010 Programme of the Ministry of Science and Innovation (MICINN, Grant No. CSD2010-00034-SimulPast) and through Grant No. CSO2012-31401 of the Ministry of Economy and Competitiveness (MINECO). Previous extended versions were presented at the Philosophy of Science Seminar of the University of Helsinki (April 2009) and at the IX Spanish Congress of Sociology (Barcelona, September 2007). I am grateful to the attendants to those sessions for their comments. The text also benefited from comments by Sandra González-Bailón, Ernest Weikert, and Luis M. Miller in several discussions in the Analytical Social Theory Seminar in Barcelona.

1. I leave aside here the case of rational choice theory, which would deserve another different chapter.
2. For instance, in the 5th edition of George Ritzer's (2000) widely used handbook of sociological theory, the terms 'reductionism' and 'reductionist' appear six times, always associated to pejorative qualifications or negative assessments of several theoretical orientations, without need to argue why 'reduction' is bad.
3. Arguably, by 'reductionism' they both mean 'elimination', which is just one of the possible meanings of the term. I will deal with this issue below.
3. See Brodbeck (1958) for a defense of this claim for the social sciences, Weinberg (1992 and 2001) for the natural sciences, and Kim (1999) or Jones (2004) for psychology and the philosophy of mind.
4. In fact, Searle identifies five different types of reduction: ontological, ontological properties reduction, theoretical, definitional, and causal. However, the ontological properties reduction, as claimed by Searle himself, is just a sub-case of ontological reduction. Here I will also deal jointly with ontological and causal reduction for reasons explained in the next section. Note also that theoretical reduction corresponds to what Udehn (2001: 324–325) calls 'scientific reduction', while definitional or linguistic reduction is equal to his 'philosophical reduction'.
5. See for instance the well-known contributions to the debates on MI by Berger and Offe (1982), or Levine, Sober and Wright (1987)
6. In more technical terms, *intensionality* is defined in the philosophy of mind and language as a property of certain sentences that do not satisfy the conditions of substitutability of coreferring expressions *salva veritate* and existential generalization (Searle, 2001: 59). Typically, sentences about beliefs, desires and volitions, which are present in many social-scientific explanations of behavior, do not satisfy those conditions, or, to say it differently, are opaque in terms of external reference.

REFERENCES

Anderson, Philip (1972) "More is different", *Science*, 177: 393–396.

Berger, Johannes and Claus Offe (1982) "Functionalism vs. rational choice? Some questions concerning the rationality of choosing one or the other", *Theory and Society*, 11(2): 521–526.

Brodbeck, May (1958) "Methodological individualisms: Definition and reduction", *Philosophy of Science*, 25(1): 1–22.

Cherkaoui, Mohamed (2005) *Invisible Codes: Essays on Generative Mechanisms*. Oxford: The Bardwell Press.

Dore, Ronald P. (1961) "Function and cause", in *The Philosophy of Social Explanation*, edited by Alan Ryan. Oxford: Oxford University Press, pp. 65–81.

Elster, Jon (1982) "Marxism, functionalism, and game theory", *Theory and Society* 11(2): 453–482.

——— (1985) *Making Sense of Marx*. Cambridge: Cambridge University Press.

——— (2007) *Explaining Social Behavior*. New York: Cambridge University Press.

Friedman, Milton (1966) "The methodology of positive economics", In *Essays in Positive*

Economics. Chicago: University of Chicago Press.
Granovetter, Mark (1978) "Threshold models of collective behavior", *American Journal of Sociology*, 83(6): 1420–1443.
Hedström, Peter, & Bearman, Peter (2009) "What is analytical sociology all about? An introductory essay", in P. Hedström & P. Bearman (eds), *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press.
Jones, Todd (2004) "Special sciences: still a flawed argument after all these years", *Cognitive Science* 28(3): 409–432.
Kim, Jaegwon (1999) "Making sense of emergence", *Philosophical Studies* 95(1–2): 3–39.
Levine, Andrew; Sober, Elliott, and Wright, Erik O. (1987) "Marxism and methodological individualism", *New Left Review*, 162: 67–84.
Little, Daniel (2012) "Explanatory autonomy and Coleman's boat", *Theoria*, 74: 137–151.
Manzo, Gianluca (2012) "Full and sketched micro-foundations: The old resurgence of a dubious distinction", *Sociologica*, 1.
Moulines, Ulises C. (2006) "Ontology, reduction, emergence: A general frame", *Synthèse* 151(3): 313–323.
Ritzer, George (2000) *Modern Sociological Theory* (5th edn.). New York: McGraw-Hill.
Searle, John R. (2001) *Rationality in Action*. Cambridge (Mass.) The MIT Press.
—— (1992) *The Rediscovery of the Mind*. Cambridge (Mass.): The MIT Press.
Weber, Max (1922) *Economía y sociedad. Esbozo de sociología comprensiva* (2 vol.) [*Economy and Society: An Outline of Interpretive Sociology*]. Mexico: Fondo de Cultura Económica, 1964.
Weinberg, Steven (1992) "Two cheers for reductionism", in *Dreams of a Final Theory*. New York: Vintage Books.
—— (2001) *Facing-Up: Science and Its Cultural Adversaries*. Harvard: Harvard University Press.
Wilson, Edward O. (1998) *Consilience. The Unity of Knowledge*, New York, Vintage Books.

11

AGGREGATION PARADOXES AND NON-COMMUTATIVITY

Dominique Raynaud
PLC, Université Pierre-Mendès-France (Grenoble)
GEMASS CNRS UMR 8598 (Paris)

INTRODUCTION

Mohammed Cherkaoui is acknowledged worldwide for his penetrating insight into the problem of transition from the micro to the macro level (see Cherkaoui, 1997; 2003). This problem has so numerous and subtle ramifications that I will proceed analytically by looking at only a very small part of it, which is the aggregation paradox.

Australian analytic philosopher Philip Pettit is known to have supported on several occasions the "common mind" thesis on the basis of the doctrinal paradox (Pettit, 1996; 2000; 2003; 2004; List and Pettit, 2002). According to him, social entities would have their own intentional states such as beliefs, desires, judgements, and intentions (Pettit, 2003: 178). As this thesis mobilizes the concept of the individual, and Pettit takes care to redefine the concepts of individualism and holism,[1] we have to start by recalling some definitions. Pettit maintains that individualism does not oppose holism, as claimed by most sociologists.

We should instead consider four related terms, generally confused with one another: individualism *vs.* atomism and holism *vs.* collectivism. Antagonism only exists between:

— *individualism vs. collectivism*, the question being whether individuals are likely to lose, or not lose, their autonomy in the interplay of historical forces.
— *holism vs. atomism*, the question being whether individuals are isolated from each other, or depend on each other, in relation to the acquisition of several basic properties such as "thinking".

Pettit explicitly defines holism as "the thesis that individuals are not fully released from any connection" (2000: 44). The negative turn of such definition makes Pettit's holism a "weak holism" compared to some versions known in sociology.

If we now compare the terms that would have been confused, we see that the holism differs from collectivism because *collectivism* builds on a vertical conception of dependence, similar to the power of coercion exerted by society on individuals, whereas *holism* introduces a horizontal conception of dependence, which is modelled on individual interdependencies that underpin social relations. Likewise, individualism is different from atomism, because *atomism* is the ability of individuals to free themselves from horizontal social ties, whereas *individualism* is the ability of individuals to escape social constraints exerted vertically on them.

The redefinition of these four terms immediately suggests the falsity of the atomistic and collectivist theses. *Collectivism* is false because individuals are always able to escape the burden of social structure. *Atomism* is false because we can never free ourselves completely from social interdependencies.

As holism and individualism are no longer conflicting, Pettit endorses a new and refined thesis, that of "individualistic holism." This thesis is indeed unproblematic, as long as the terms holism and individualism are used according to the author's definitions.

THE DOCTRINAL PARADOX

"Groups with Minds of Their Own" (Pettit, 2003: 167–193; 2004: 129–172) is probably the text in which Pettit's holistic thesis is the most advanced, because the irreducibility of the whole to the parts is here firmly supported, as social groups are seen as "collective intentional subjects". The outline of Pettit's paper is the following:

1. He presents the doctrinal paradox, which first appeared in jurisprudence (Kornhauser and Sager, 1993; 2004);
2. He describes a generalized version of it, called discursive dilemma, because there is a need to encompass any conjunctive or disjunctive decision made by any group of people;
3. Building on the discursive dilemma, he finally argues that "social integrates are intentional subjects" (Pettit, 2003: 178).

Therefore (3) depends on (2), which in turn depends on (1). The whole argument is built on the doctrinal paradox—whatever its name—to which most of this chapter will therefore be dedicated.

Let us explain what the doctrinal paradox is with an example. A court, consisting of three judges, must decide whether a person is liable or not for breaking a contract with the plaintiff (Kornhauser and Sager, 1993: 11; Pettit, 2003: 168). The decision of the judges, based on the examination of premises, has a binary form: each issue can only be accepted (1) or rejected (0). MAJ denotes the majority opinion of the court on any issue relevant to the judgement.

	Contract?	*Breach?*	*Liable?*
A	1	0	0
B	0	1	0
C	1	1	1
MAJ	1	1	1/0

Matrix 1

The decision of the court will vary according to the order of the steps chosen by the judges, who can aggregate either their premises (premise-based procedure) or their conclusions (conclusion-based procedure).

Premise-based procedure. Is there a contract? Yes, because MAJ(1, 0, 1) = 1. Is there a breach in the contract? Yes, because MAJ(0, 1, 1) = 1. Now, since AND(1, 1) = 1, the person is liable for breaking the contract. *Conclusion-based procedure*. Is the person liable? No, because MAJ(0, 0, 1) = 0. We are facing a situation where the same bottom right cell has a different number according to the procedure chosen by the judges. As the content differs, a distinction should be made between the *individual decision* of the judges and the *collective decision* of the court. The "court" is endowed with intentional states of its own, irreducible to those of the judges as individuals. This conclusion is crystal-clear.

The doctrinal paradox is akin to a broad range of aggregative paradoxes such as the Condorcet paradox, the Ostrogorski paradox or the multiple election paradox within the literature on Social Choice Theory. Condorcet's result can be expressed as follows: Given a set of transitive individual preferences (complete strict order relations), the aggregation does not necessarily retain transitivity of the individual preferences. Cases of intransitivity have a maximum probability $p = 0.0877$ for three candidates and an infinite population. The loss of transitivity of the individual preferences has been discussed at length in the literature (Borda, 1770; Condorcet, 1785; Laplace, 1812; Arrow, 1951; Guilbaud, 1952; Granger, 1956; Sen, 1970; Mimiague and Rousseau, 1973; Rashed, 1974; Arkhipoff, 1985; 1986; 1988). Pioneering works by Borda and Condorcet are as much ignored as echoed in sociological theory of aggregation (Olson, 1965; Schelling, 1971; 1978; Alexander, 1987 *vs*. Coleman, 1964; 1990; Boudon, 1997; 1998; Cherkaoui, 1997; 2003).

We owe to Oleg Arkhipoff (1988) a drastic simplification of the data used by Condorcet to illustrate the loss of transitivity. There is no need to assume sophisticated conditions[2] to observe the loss of transitivity at the collective level. It already appears on the "reduced case" with three candidates *A, B,* and *C*, and three voters *P, Q,* and *R* (Arkhipoff, 1988: 265).

	$A \succ B$?	$B \succ C$?	$C \succ A$?
P	1	1	0
Q	0	1	1
R	1	0	1
MAJ	1	1	1

Matrix 2

Although preferences are transitive at the individual level (there is one 0 in each line), they are not at the collective level: MAJ(1, 0, 1) = 1, MAJ(1, 1, 0) = 1, MAJ(0, 1, 1) = 1. This means that *A* is preferred to *B*, *B* is preferred to *C*, and *C* is preferred to *A*; thus no candidate is elected. The doctrinal paradox and the reduced case of the Condorcet paradox are akin. Discussion can be engaged on this basis.

I. IS THE DOCTRINAL PARADOX A PARADOX?

That doctrinal paradox is not a paradox can be established in two different ways, depending on whether the doctrinal paradox is equated, or not, to Condorcet paradox or other aggregation paradoxes.

Case 1. Consider first that the doctrinal and Condorcet paradoxes are reducible to each other, as has been suggested by several authors (List and Pettit, 2002; Pigozzi, 2005; Bovens and Rabinowicz, 2006). Oleg Arkhipoff (1988) has convincingly shown that all aggregative paradoxes, of which the "Condorcet paradox" is the best known example, are not paradoxes. His main argument is that a "collective preference" is not a preference, exactly as the "average man" is not a man. We can go further and show why this is a simple case of homonymy. Let us compare the ways in which individual and collective preferences are formed within the Condorcet theorem. Individual preference is left to the discretion of the voter. Collective preference results from the sum of the votes under the rule of majority. The collective rule has no equivalent—not even

on pro rata—in the formation of individual preferences. As construction rules are different, individual and collective preference are not of the same nature. The use of the same word is a mistake. Cases falling under the Condorcet theorem give rise to endless glosses only when one confuses, by an artifice of language, facts that occur at the individual level and facts that occur at the collective level. Arkhipoff is perfectly correct when he says that "every aggregative paradox is a fallacy of composition", that "the common nature of aggregative paradoxes is that the aggregated object is from another species than that of the objects aggregated", and that "there is a change in meaning, when one moves from one level of aggregation to another" (Arkhipoff, 1988: 254).

If the doctrinal paradox and Condorcet paradox are well within the same class of aggregative paradoxes, and if aggregative paradoxes are not paradoxes, then the doctrinal paradox is not a paradox either. We can ensure this, by studying the construction of individual and collective choices in the example of the three judges. The choice of a judge is based on consideration of two premises: is there a contract? is there a breach? The collective choice of the court results from a vote by majority: the individual will be held liable if at least two out of three judges state that he broke the contract. So it comes as no surprise that the individual and collective choice are different: they are not built in the same manner. What is salient in the "collective choice" is not choice, but *collective aggregation*. The route we have taken leads to a first criticism of the doctrinal paradox, which is a semantic criticism.

Case 2. Let us now assume that the doctrinal paradox is other than Condorcet paradox. In what way is it different? A distinctive feature of the doctrinal paradox is that the result of aggregation varies according to the order of the steps taken by the three judges.

1. Let us generalize the way to denote individuals and opinions in any social group. Consider a social group $G_N = \{A, B, C \dots J \dots N\}$, each member of which has a set of m opinions $O_m = \{a_J, b_J, c_J \dots i_J \dots m_J\}$. Any elementary opinion of an individual J is denoted i_J. Next two functions are defined.

The first one is the *deduction* $\gamma(i_J)$ that any individual can make from his opinions. This function is defined as

$$\gamma : \{1, 0\}^m \to \{1, 0\}, (a_J, b_J, c_J \dots i_J \dots m_J) \longmapsto \gamma(i_J)$$

Deduction transforms a line-vector into a number (or, equivalently, a matrix into a column-vector). As seen by List and Pettit (2002), the truth table of deduction is equal to that of the conjunction: $\gamma(i_J) = \text{AND}(a_J, b_J, c_J \dots i_J \dots m_J)$.

The second function is *aggregation* $\alpha(i_J)$ that the group G makes on any issue i. This function is defined as

$$\alpha : \{1, 0\}^N \to \{1, 0\}, (i_A, i_B, i_C \dots i_J \dots i_N) \longmapsto \alpha(i_J)$$

Aggregation transforms a column-vector into a number (or a matrix into a line-vector). Aggregation has the truth table of the function majority: $\alpha(i_J) = \text{MAJ}(i_A, i_B, i_C \dots i_J \dots i_N)$.

The matrix of opinions i_J with the two functions $\gamma(i_J)$ and $\alpha(i_J)$ appears as:

	a	b		i		m	$\gamma(i_J)$
A	a_A	b_A	…	i_A	…	m_A	$\gamma(i_A)$
B	a_B	b_B	…	i_B	…	m_B	$\gamma(i_B)$
	…	…	…	…	…	…	
J	a_J	b_J	…	i_J	…	m_J	$\gamma(i_J)$
	…	…	…	…	…	…	
N	a_N	b_J	…	i_N	…	m_N	$\gamma(i_N)$
$\alpha(i_J)$	$\alpha(a_J)$	$\alpha(b_J)$		$\alpha(i_J)$		$\alpha(m_J)$	?

Matrix 3

2. Once this notation is adopted, it is easy to see that the paradox lies in the order of the steps chosen by the judges. As the paradox arises from the mixing of unrelated procedures, it is good practice to show the two diagrams separately:

	a	b	$\gamma(i_J)$
A	1	0	
B	0	1	
C	1	1	
$\alpha(i_J)$	1	1	1

Matrix 4A

	a	b	$\gamma(i_J)$
A	1	0	0
B	0	1	0
C	1	1	1
$\alpha(i_J)$			0

Matrix 4B

Matrix 4A shows the premise-based procedure. If the judges aggregate their premises before deducing the judgement, they reach the conclusion that $\alpha\gamma(i_J) = 1$ (liable). Matrix 4B shows the conclusion-based procedure. If the judges draw their own conclusions before aggregating them, the outcome is $\gamma\alpha(i_J) = 0$ (not liable). The judgement is thus different when the aggregation is made before or after the deduction. The doctrinal paradox reduces to:

$$\begin{array}{ccc} i_J & \xrightarrow{\gamma} & \gamma(i_J) \\ & & \downarrow{\scriptstyle \alpha} \\ & & \alpha\gamma(i_J) \end{array} \quad \neq \quad \begin{array}{ccc} i_J & & \\ \downarrow{\scriptstyle \alpha} & & \\ \alpha(i_J) & \xrightarrow{\gamma} & \gamma\alpha(i_J) \end{array}$$

or, more simply:

$$\alpha\gamma(i_J) \neq \gamma\alpha(i_J) \tag{1}$$

This seems to be an example of non-commutativity of function composition. However this result should be proven in general, that is, for any values of the opinion matrix and for any type of deduction. We will consider in turn deduction by conjunction (Section 3), disjunction (Section 4), and majority (Section 5).

3. As conjunction and majority are logical functions, they can be written according to their associated Boolean expressions:

$$\mathrm{AND}(x, y) = xy \tag{2}$$

$$\mathrm{MAJ}(x, y, z) = yz + xz + xy \tag{3}$$

Let us rewrite the two procedures.

3.1. Deduction then Aggregation. Applying (eq. 2) horizontally:

$$\gamma(i_A) = (a_A b_A)$$
$$\gamma(i_B) = (a_B b_B)$$
$$\gamma(i_C) = (a_C b_C)$$

Applying (eq. 3) vertically:[3]

$$\alpha\gamma(i_J) = \underline{a_B b_B a_C b_C} + \underline{a_A b_A a_C b_C} + \underline{a_A b_A a_B b_B}$$

3.2. Aggregation then Deduction. Applying (eq. 3) vertically:

$$\alpha(a_J) = a_B a_C + a_A a_C + a_A a_B$$
$$\alpha(b_J) = b_B b_C + b_A b_C + b_A b_B$$

Applying (eq. 2) horizontally:

$$\gamma\alpha(i_J) = (a_B a_C + a_A a_C + a_A a_B)(b_B b_C + b_A b_C + b_A b_B)$$

Developing:

$$\begin{aligned}\gamma\alpha(i_J) = \underline{a_B a_C b_B b_C} &+ a_B a_C b_A b_C + a_B a_C b_A b_B \\ &+ a_A a_C b_B b_C + \underline{a_A a_C b_A b_C} + a_A a_C b_A b_B \\ &+ a_A a_B b_B b_C + a_A a_B b_A b_C + \underline{a_A a_B b_A b_B}\end{aligned}$$

3.3. Conclusion. Comparing $\alpha\gamma(i_J)$ and $\gamma\alpha(i_J)$ term by term, it appears that there are three identical (underlined) terms in the two developments. We can then write:

$$\gamma\alpha(i_J) = \alpha\gamma(i_J) + (a_B a_C b_A b_C + \ldots + a_A a_B b_A b_C) \text{ (six terms)}$$

Thus in general:

$$\alpha\gamma(i_J) \neq \gamma\alpha(i_J) \qquad (4)$$

The composition of functions MAJ(i_J) and AND(i_J) is non-commutative.

4. Other cases of doctrinal paradox involve logical disjunction. Suppose now three judges have to decide whether someone should be given a new trial, on the basis of two premises: inadmissible evidence or forced confession (Kornhauser and Sager, 1993: 15; Pettit, 2003: 170).

Let us write disjunction and majority functions as Boolean expressions:

$$\mathrm{OR}(x, y) = x + y \tag{5}$$

$$\mathrm{MAJ}(x, y, z) = yz + xz + xy \tag{6}$$

Now rewrite each of the two procedures.

4.1. Deduction then Aggregation. Applying (eq. 5) horizontally:

$\gamma(i_A) = a_A + b_A$
$\gamma(i_B) = a_B + b_B$
$\gamma(i_C) = a_C + b_C$

Applying (eq. 6) vertically:

$$\alpha\gamma(i_J) = (a_B + b_B)(a_C + b_C) + (a_A + b_A)(a_C + b_C) + (a_A + b_A)(a_B + b_B)$$

Developing[4]:

$$\begin{aligned}\alpha\gamma(i_J) = \underline{a_B a_C} + a_B b_C + b_B a_C + \underline{b_B b_C} \\ + \underline{a_A a_C} + a_A b_C + b_A a_C + \underline{b_A b_C} \\ + \underline{a_A a_B} + a_A b_B + b_A a_B + \underline{b_A b_B}\end{aligned}$$

4.2. Aggregation then Deduction. Applying (eq. 6) vertically:

$\alpha(a_J) = a_B a_C + a_A a_C + a_A a_B$
$\alpha(b_J) = b_B b_C + b_A b_C + b_A b_B$

Applying (eq. 5) horizontally:

$$\gamma\alpha(i_J) = \underline{a_B a_C} + \underline{a_A a_C} + \underline{a_A a_B} + \underline{b_B b_C} + \underline{b_A b_C} + \underline{b_A b_B}$$

4.3. Conclusion. Comparing $\alpha\gamma(i_J)$ and $\gamma\alpha(i_J)$ term by term, it appears that there are six identical (underlined) terms in the two developments. We can then write:

$\alpha\gamma(i_J) = \gamma\alpha(i_J) + (a_B b_C + \ldots + b_A a_B)$ (six terms)

Therefore in general:

$$\alpha\gamma(i_J) \neq \gamma\alpha(i_J) \qquad (7)$$

The composition of functions MAJ(i_J) and OR(i_J) is non-commutative.

5. Consider a variant of the Ostrogorski paradox, in which there is a two-party system and three voters having to choose the candidate with whom they agree in majority on three issues (Ostrogorski 1902; Pigozzi 2005). As functions $\alpha(i_J)$ and $\gamma(i_J)$ are majority functions, we are faced with an iteration MAJ(MAJ(i_J)).

	a	b	c	$\gamma(i_J)$
A	1	1	0	
B	1	0	0	
C	0	1	0	
$\alpha(i_J)$	1	1	0	1

Matrix 5A

	a	b	c	$\gamma(i_J)$
A	1	1	0	1
B	1	0	0	0
C	0	1	0	0
$\alpha(i_J)$				0

Matrix 5B

Matrix 5A shows the premise-based procedure. If aggregation is applied first, MAJ(1, 1, 0) = 1, MAJ(1, 0, 1) = 1, MAJ(0, 0, 0) = 0 and MAJ(1, 1, 0) = 1. Thus $\gamma\alpha(i_J) = 1$. Matrix 5B shows the conclusion-based procedure. If personal deduction is made first, MAJ(1, 1, 0) = 1, MAJ(1, 0, 0) = 0, MAJ (0, 1, 0) = 0, MAJ(1, 0, 0) = 0. Thus $\alpha\gamma(i_J) = 0$.

In this example, the result is sensitive to the order of the steps. This suggests that composition of majority functions is also non-commutative, a result that may be proven by using the Boolean expression:

$$\mathrm{MAJ}(x, y, z) = yz + xz + xy \tag{8}$$

5.1. Deduction then Aggregation. In this case, the two successive majority functions γ and α operate in that way:

$$\begin{pmatrix} a_A & b_A & c_A \\ a_B & b_B & c_B \\ a_C & b_C & c_C \end{pmatrix} \xrightarrow{\gamma} \begin{pmatrix} \gamma(i_A) \\ \gamma(i_B) \\ \gamma(i_C) \end{pmatrix} \xrightarrow{\alpha} \alpha\gamma(i_J)$$

Applying (eq. 8) horizontally:

$\gamma(i_A) = (b_Ac_A + a_Ac_A + a_Ab_A)$
$\gamma(i_B) = (b_Bc_B + a_Bc_B + a_Bb_B)$
$\gamma(i_C) = (b_Cc_C + a_Cc_C + a_Cb_C)$

Applying now (eq. 8) vertically:

$$\begin{aligned}\alpha\gamma(i_J) = {} & (b_Bc_B + a_Bc_B + a_Bb_B)(b_Cc_C + a_Cc_C + a_Cb_C) \\ & + (b_Ac_A + a_Ac_A + a_Ab_A)(b_Cc_C + a_Cc_C + a_Cb_C) \\ & + (b_Ac_A + a_Ac_A + a_Ab_A)(b_Cc_C + a_Cc_C + a_Cb_C)\end{aligned}$$

Developing[5]:

$$\begin{aligned}\alpha\gamma(i_J) = {} & (\underline{b_Bc_Bb_Cc_C} + b_Bc_Ba_Cc_C + b_Bc_Ba_Cb_C + a_Bc_Bb_Cc_C + \underline{a_Bc_Ba_Cc_C} + \\ & a_Bc_Ba_Cb_C + a_Bb_Bb_Cc_C + a_Bb_Ba_Cc_C + \underline{a_Bb_Ba_Cb_C} + \underline{b_Ac_Ab_Cc_C} + b_Ac_Aa_Cc_C + \\ & b_Ac_Aa_Cb_C + a_Ac_Ab_Cc_C + \underline{a_Ac_Aa_Cc_C} + a_Ac_Aa_Cb_C + a_Ab_Ab_Cc_C + a_Ab_Aa_Cc_C + \\ & \underline{a_Ab_Aa_Cb_C} + \underline{b_Ac_Ab_Cc_C} + b_Ac_Aa_Cc_C + b_Ac_Aa_Cb_C + a_Ac_Ab_Cc_C + \underline{a_Ac_Aa_Cc_C} + \\ & a_Ac_Aa_Cb_C + a_Ab_Ab_Cc_C + a_Ab_Aa_Cc_C + \underline{a_Ab_Aa_Cb_C})\end{aligned}$$

5.2. Aggregation then Deduction. In that case, the two successive majority functions α and γ operate in that way:

$$\begin{pmatrix} a_A & b_A & c_A \\ a_B & b_B & c_B \\ a_C & b_C & c_C \end{pmatrix} \xrightarrow{\alpha} \begin{pmatrix} \alpha(a_J) \\ \alpha(b_J) \\ \alpha(c_J) \end{pmatrix} \xrightarrow{\gamma} \gamma\alpha(i_J)$$

Applying (eq. 8) vertically:

$\alpha(a_J) = (a_B a_C + a_A a_C + a_A a_B)$
$\alpha(b_J) = (b_B b_C + b_A b_C + b_A b_B)$
$\alpha(c_J) = (c_B c_C + c_A c_C + c_A c_B)$

Applying now (eq. 8) horizontally:

$$\gamma\alpha(i_J) = (b_B b_C + b_A b_C + b_A b_B)(c_B c_C + c_A c_C + c_A c_B) + (a_B a_C + a_A a_C + a_A a_B)\,(c_B c_C + c_A c_C + c_A c_B) + (a_B a_C + a_A a_C + a_A a_B)\,(b_B b_C + b_A b_C + b_A b_B)$$

Developing:

$$\gamma\alpha(i_J) = (\underline{b_B b_C c_B c_C} + b_B b_C c_A c_C + b_B b_C c_A c_B + b_A b_C c_B c_C + \underline{b_A b_C c_A c_C} + b_A b_C c_A c_B + b_A b_B c_B c_C + b_A b_B c_A c_C + \underline{b_A b_B c_A c_B} + \underline{a_B a_C c_B c_C} + a_B a_C c_A c_C + a_B a_C c_A c_B + a_A a_C c_B c_C + \underline{a_A a_C c_A c_C} + a_A a_C c_A c_B + a_A a_B c_B c_C + a_A a_B c_A c_C + \underline{a_A a_B c_A c_B} + \underline{a_B a_C b_B b_C} + a_B a_C b_A b_C + a_B a_C b_A b_B + a_A a_C b_B b_C + \underline{a_A a_C b_A b_C} + a_A a_C b_A b_B + a_A a_B b_B b_C + a_A a_B b_A b_C + \underline{a_A a_B b_A b_B})$$

5.3. Conclusion. Nine identical (underlined) terms are found in $\alpha\gamma(i_J)$ and $\gamma\alpha(i_J)$. If the sum of these terms is denoted by A, the results of the two procedures can be written as:

$\alpha\gamma(i_J) = A + (b_B c_B a_C c_C + \ldots + a_A b_A a_C c_C)$ (18 terms)
$\gamma\alpha(i_J) = A + (b_B b_C c_A c_C + \ldots + a_A a_B b_A b_C)$ (18 terms)

No further simplification is possible, because the remaining terms are different. There is no chance in general for the two results to be equal. We thus draw the conclusion:

$$\alpha\gamma(i_J) \neq \gamma\alpha(i_J) \tag{9}$$

The iteration MAJ(MAJ(i_J)) is non-commutative.

6. The outcome of Sections 3 to 5 is that procedures $\alpha\gamma$ and $\gamma\alpha$ do not commute in general. They only do so, if the opinion matrix is square

and symmetrical ($i_J = j_I$), a condition that is hard to fulfil because I, J are indexing individuals, and i, j are indexing issues brought to the debate.

Another interesting point is that, as the iterations of functions AND(i_J) and OR(i_J) are commutative, the asymmetry at the root of the doctrinal paradox is only due to the function MAJ(i_J), when combined with other functions. Hence the conclusion:

The doctrinal paradox (discursive dilemma, and presumably most of aggregation paradoxes) arises from the non-commutativity of function composition, whatever the functions composed, provided that one of them is the majority function.

However, it is worth noting that non-commutativity[6] is nothing special in itself, because the composition of functions is generally non-commutative: $g \circ f\ (x) \neq f \circ g\ (x)$. A wide range of examples is known in science: multiplication of matrices, vector product, quaternion product, non-commutative geometry, set theory, groups of symmetry (dihedral groups) … Non-commutativity is also very common in the real world: kinship systems, displacements in space by translation and rotation, everyday activities such as washing and drying, eating and doing the dishes …, and so on.

Since non-commutative operations and activities are not recognized as paradoxes, there is no reason to consider the doctrinal paradox as a paradox either. This is a different semantic criticism, namely, the criticism of the propensity to identify paradoxes where there are none.

II. DOES THE DOCTRINAL PARADOX WORK?

The doctrinal paradox supposes the premises to be separable. This is often made explicit: "There is a conclusion to be decided among the judges by reference to a conjunction of separable premises" (Pettit, 2003: 192). Nonetheless separability is hardly ensured in the examples cited.

Case 1. Let us return to the example of the three judges who must decide whether a person is liable for breaking a contract or not (Pettit 2003, 169). Although the premises are explicitly requested to be separable, it turns out that they are not in this case.

Matrix 6A presents the premise-based procedure; matrix 6B presents the conclusion-based procedure. The doctrinal paradox is due to $\gamma\alpha(i_J) = 1$ whereas $\alpha\gamma(i_J) = 0$. Look deeper into the data. Judge B claims that the contract has been broken ($b_B = 1$) but however contends that the contract does not exist ($a_B = 0$). These views are mutually exclusive. Any breach of contract presupposes the existence of the contract. This internal flaw requires correction of the data (cancelled values).

	a	b	$\gamma(i_J)$
A	1	0	
B	0	~~1~~	
C	1	1	
$\alpha(i_J)$	1	~~1~~	~~1~~

Matrix 6A

	a	b	$\gamma(i_J)$
A	1	0	0
B	0	~~1~~	0
C	1	1	1
$\alpha(i_J)$			0

Matrix 6B

If Judge B says there is no contract, he cannot claim that the contract has been broken, so $b_B = 0$. Let us correct the numbers according to the premise-based procedure: $\alpha(b_J) = 0$, for MAJ(0, 0, 1) = 0. As a result: $\gamma(\alpha(i_J)) = 0$, for AND(1, 0) = 0. Let us now correct the numbers in the conclusion-based procedure: $\gamma(i_B) = 0$, because AND(0, 0) = 0. Thus $\alpha\gamma(i_J) = 0$, because MAJ(0, 0, 1) = 0. We reach the conclusion that $\gamma\alpha(i_J) = \alpha\gamma(i_J)$. Therefore the result is the same, regardless of the procedure. The paradox collapses when the data are correct.

Case 2. Pettit (2003) reviews five cases of doctrinal paradox. If we discard the example previously discussed (Case 1), there are four examples. In Matrix 1, premises are not separable because an individual can not be found liable of negligence when he has no duty to help others (2003: 168). In Matrix 4, premises are not separable because an individual does not have to implement a specific measure to curb a danger when such danger is not proven (2003: 171). In Matrix 5, premises are not separable because one cannot decide to increase defence spending without increasing taxes or reducing other spending (2003: 178). Matrix 3 offers the only case where the premises are separable, probably because each judge's conclusion is to be made on the basis of a logical disjunction (Part I, Section 4).

In most cases, the doctrinal paradox collapses and its consequences are alike. In four cases out of five, the decision of the court does not oppose the individual conclusion of each judge. Theses cases simply show that the result is insensitive to the premise-based or conclusion-based procedures or, equivalently, that aggregation is consistent with the opinion of the majority. In only one case out of five does the doctrinal paradox resist criticism; however this case is still undergoing the criticism that non-commutativity is nothing special.

CONCLUSION

This chapter's aim was intended to examine the doctrinal paradox, by both external criticism of its significance (Part I) and internal criticism of its contents (Part II). We have reached the following results:

Part I: The doctrinal paradox is not a paradox,
Case 1: whether reduced to the Condorcet paradox,
Case 2: or not.
Part II: The doctrinal paradox has a fatal flaw, which is the non-separability of the premises,
Case 1: in the oft-told example of the contract broken,
Case 2: as in other cases.

If we consider that one case resists criticism of Part II, *the main result of this study is reinterpreting aggregative paradoxes as the non-commutativity of function composition*. This also suggests avoiding the name "paradox" nonetheless abundantly used in the literature on Social Choice Theory. Voting paradoxes are often said "unpleasant surprises", "bizarre" or "unfair outcomes" (Nurmi, 1999: vii). A first step was taken by renaming the doctrinal paradox the discursive dilemma, because "the problem represents a hard choice or dilemma, not anything that strictly deserves to be called a paradox" (Pettit, 2003: 174). We could go further by noting that the word "paradox" has always a subjective foundation. For whom, and in which context, do such examples turn out to be unpleasant or bizarre? This is undefined. *Instead, non-commutativity is a universal and clearly defined concept.*

The non-separability of premises described in Part II suggests that, in most cases, the doctrinal paradox (1) and the discursive dilemma alike (2) cannot serve as a basis to the thesis that groups have minds of their own (3). This result undermines the conclusion that "Any groups that seek to make deliberative, reasoned judgments, then, face a dilemma. They may maximize responsiveness to individual views, running the risk of collectively endorsing inconsistent sets of propositions. Or they may can impose the discipline of reason at the collective level, running the risk of collectively endorsing a conclusion that a majority of them—perhaps even all of them—individually reject" (Pettit, 2003: 175).

This does not mean that new ideas about individualism and holism are necessarily doomed to failure—some convincing attempts are known (Bunge, 1979; 2004). This rather means that the doctrinal paradox is not very good at supporting this thesis.

NOTES

1. These theses shall be understood in the ontological sense. As ontology does not necessarily interfere with methodology, ontological collectivism is not incompatible with methodological individualism; ontological holism is not incompatible with methodological atomism; ontological individualism is not incompatible with methodological collectivism; ontological atomism is not incompatible with methodological holism.
2. For instance, in Borda's memoir, the electoral body was composed of 21 voters who had to vote among three candidates. The distribution of preferences needed for the loss of transitivity was 1 ($A \succ B$, $B \succ C$), 7 ($A \succ C$, $C \succ B$), 7 ($B \succ C$, $C \succ A$), and 6 ($C \succ B$, $B \succ A$) (Borda 1781, 660). In Condorcet's memoir, the electoral body was composed of 60 voters and had to vote among three candidates. The distribution of preferences needed for the loss of transitivity was 23 ($A \succ B$, $B \succ C$), 17 ($B \succ C$, $C \succ A$), 2 ($B \succ A$, $A \succ C$), 10 ($C \succ A$, $A \succ B$), and 8 ($C \succ B$, $B \succ A$) (Condorcet 1785, lviii).
3. Identical terms in both developments are underlined.
4. Identical terms in both developments are underlined.
5. Identical terms in both developments are underlined.
6. Today, the word "commutative" is almost exclusively used to describe operations, in which the operands can be reversed with no change in result (addition, multiplication *vs.* subtraction, division). However, the term "commutative" was first introduced by Servois (1814, 98) to describe composition of functions.

REFERENCES

Alexander, J.C. (1987) *The Micro–Macro Link*. Berkeley, CA: University of California Press.

Arkhipoff, O. (1985) "Pour une théorie générale de l'agrégation I", *Journal de la Société Statistique de Paris*, 126(4): 145–167.

——— (1986) "Pour une théorie générale de l'agrégation II", *Journal de la Société Statistique de Paris*, 127(1): 27–43.

——— (1988) "La fin du paradoxe de Condorcet. Quelles conséquences pour la comptabilité nationale?" *Journal de la Société Statistique de Paris*, 129(4): 244–268.

Arrow, K.J. (1951) *Social Choice and Individual Values*. New York: Wiley.

Borda, J.-C. de (1770) "Mémoire sur les élections au scrutin", *Histoire de l'Académie Royale des Sciences*. Paris: Imprimerie Royale, 1781, pp. 657–665.

Boudon, R. (1997) "Le paradoxe du vote et la théorie de la rationalité", *Revue Française de Sociologie*, 38(2): 217–227.

——— (1998) "Social mechanism without black boxes", in Hedström, P. and Swedberg, R. (eds), *Social Mechanisms. An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press, pp. 172–203.

Bovens, L. and Rabinowicz, W. (2006) "Democratic answers to complex questions. An epistemic perspective", *Synthese*, 150(1): 131–153.

Bunge, M. (1979) "A system concept of society. Beyond individualism and holism", *Theory and Decision* 10: 13–30.

——— (2004) "How does it work? The search for explanatory mechanisms", *Philosophy of the Social Sciences,* 34(2): 182–210.

Cherkaoui, M. (1997) "Le réel et ses niveaux. Peut-on toujours fonder la macrologie sur la micrologie?" *Revue Française de Sociologie*, 38(3): 497–524.

——— (2003) "Les transitions micro–macro. Limites de la théorie du choix rationnel dans les *Foundations of Social Theory*", *Revue Française de Sociologie*, 44(2): 231–254.

Coleman, J. (1964) *Introduction to Mathematical Sociology*. New York: The Free Press.

——— (1990) *Foundations of Social Theory*. Cambridge, MA: Belknap Press of Harvard University.

Condorcet, J.A.N. Caritat de (1785) *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris: Imprimerie Royale.

Granger, G.-G. (1956) *La mathématique sociale du Marquis de Condorcet*. Paris: PUF.

Guilbaud, G.-T. (1952) "Les théories de l'intérêt général et le problème logique de l'agrégation", *Economie Appliquée*, 5(4): 501–584.

Kornhauser, L.A. and Sager, L.G. (1993) "The one and the many: Adjudication in collegial courts", *California Law Review*, 91: 1–51.

——— (2004) "Group choice in paradoxical cases", *Philosophy and Public Affairs*, 32: 249–276.

Laplace, P.S. de (1812) "Application du calcul des probabilités aux sciences morales", *Théorie analytique des probabilités*. *Oeuvres complètes de Laplace*, published under the auspices of the Académie des Sciences. Paris: Gauthiers-Villars, 1886, pp. lxxviii–cxii.

List, C., Pettit, P. (2002) "Aggregating sets of judgments: An impossibility result", *Economics and Philosophy*, 18: 89–110.

Mimiague, F. Rousseau, J.M. (1973) "Effet Condorcet: typologie et calcul de fréquences", *Mathématiques et Sciences Humaines*, 43: 7–27.

Nurmi, H. (1999) *Voting Paradoxes and How to Deal with Them*. Berlin: Springer.

Olson, M. (1965) *The Logic of Collective Action*. Cambridge, MA: Harvard University Press.

Ostrogorski, M.Y. (1902) *La Démocratie et l'organisation des partis politiques*. Paris: Calmann-Lévy.

Pettit, P. (1996) *The Common Mind*. Oxford: Oxford University Press.

——— (2000) "Défense et définition du holisme social", in P. Livet and R. Ogien (eds), *L'Enquête ontologique. Du mode d'existence des objets sociaux*. Paris: Éditions de l'EHESS, pp. 43–63.

——— (2003) "Groups with minds of their own", in F.F. Schmitt (ed.), *Socializing Metaphysics. The Nature of Social Reality*. Oxford: Rowman & Littlefield Publishers, pp. 167–193.

——— (2004) *Penser en société. Essais de métaphysique sociale et de méthodologie*. Paris: PUF.

Pigozzi, G. (2005) "Two aggregation paradoxes in social decision making: The Ostrogorski paradox and the discursive dilemma", *Episteme: A Journal of Social Epistemology*, 2(2): 22–42.

Rashed, R. (1974) *Condorcet. Mathématique et société*. Paris: Hermann.

Schelling, T.C. (1971) "Dynamic models of segregation", *Journal of Mathematical Sociology*, 1: 143–186.

——— (1978) *Micromotives and Macrobehavior*. New York: W.W. Norton and Co.

Sen, A.K. (1970) *Collective Choice and Social Welfare*. London: Oliver and Boyd.

Servois, François Joseph (1814–5) "Analise transcendante. Essai sur un nouveau mode d'exposition des principes du calcul différentiel", *Annales de Mathématiques Pures et Appliquées*, 5: 93–140.

PART THREE

Social Theory

12

RITUALS AS MECHANISMS

Filippo Barbera
University of Turin and Collegio Carlo Alberto—Moncalieri

Nicola Negri
University of Turin

INTRODUCTION

Supporters of mechanisms-based explanation (Coleman, 1987; Goldthorpe, 2000), criticize Durkheim for his positivist conception of science. Mohamed Cherkaoui notes that this positivist perspective stresses a radical refusal of metaphysics and therefore any ontological consideration about "deep" causation processes embedded in different levels of reality (Cherkaoui, 2005: 78–80 and 2000). In classical positivism, the argument goes on, this (op)position leads to deny any argument that directly looks at "how phenomena are produced", in favour of the idea of causality as a matter of concomitant variations. Durkheim, according to the received knowledge of the positivist conception of science, has proposed an explanatory model which take little, if no account at all of the roles of reality levels (*micro–meso–macro*), thus undervaluing the role of interaction and interdependence of actions and therefore only analysing macro-level phenomena as the

sum of independent actions (*ibidem*). With regard to this criticism, on the other hand, Raymond Boudon (1993) and Cherkaoui (2000; 2008) interpretations point to the role of *meso* and *micro*-level mechanisms in Durkheim's work (Boudon 1993).

In following this path we will first of all look at Durkheim's explanatory model, taking into account a "theoretical asymmetry" in the analysis of macro–micro and micro–macro mechanisms, as outlined by Cherkaoui (2008: ch. 1). In the second section, the attempt to rebuild this asymmetry will leads us to the macro–micro *immergentist*[1] mechanisms that support the genesis of social identities, recognition patterns and comparability schemes in social action. In the third section, these mechanisms will be conceptualized as ritual processes. In the fourth section, we will look at the connection between the theory of networks and that of rituals, within the action's comparability schemes that Harrison White calls "disciplines" (1992; 2001). Finally, in the fifth and concluding sections, we will look at the implications of ritual's theory for the conceptualization of social change.

This paper is hence providing a starting point to fill the gap between analytical sociology and the so-called "thick description" of the cultural meaning of actions (Geertz 2000), which seem to have more which divide them than unite them (Gross 2009; see Manzo, 2010 for a reply). For instance, Peter Hedström and Lars Udehn (2009: 29) distinguish analytical sociology from thick description as much for the lesser degree of inclusiveness of the former (the analytical isolation of certain dimensions and mechanisms of *explanans* at the expense of others) as much as for its minor generality (the extension of the *explanandum*).

Nonetheless, the meaning of social action and its cognate concepts are not alien to an analytical sociology framework, rather they are sometimes boiled down to other concepts, such as cognitive rationality (Boudon, 1998), framing (Lindenberg, 2000; Esser, 1993), signals (Gambetta, 2009), beliefs (Rydgren, 2009) and still others (Manzo, 2010: 150–157). On this ground, we argue, the explanation of the *meaning* of social action can be included in the macro–micro–macro framework, even with regard to its most complex implications such as social identity and non-intentional processes. This approach would confirm the purely methodological nature of analytical sociological schemes, which

we believe is compatible with a wide spectrum of substantive theories (for a statement, see Manzo 2010 and 2012).

1. DURKHEIMIAN MECHANISMS: A DEBATE

According to James S. Coleman the micro–macro connection in Durkheim's "Suicide" boils down to an additive process of individual and independent actions:

> In *Suicide,* Durkheim (….) explains the difference between the rate of suicide in various societies and between various subgroups within these societies. Even given his insistence that suicide is a social act and his argument against social psychology, even Durkheim explains it in terms of individual behavior (Coleman, 1987: 153).

At best, as Cherkaoui pointed out: "(…) Those who commit suicide (…) constitute a coherent group inasmuch they commit their act in conformity with an external cause" (2008: 24). This means that the observed behavioural uniformities that makes all actors adjust their actions in a similar manner are due to an environmental effect (Hedström, 2005), leaving the micro to macro link unexplained and not qualifying suicide as an emergent phenomena.

These options mark out a "Durkheimian" approach to suicide research, namely: "an approach that focuses on associations between suicide rates in different population groups and various properties of these groups" (Hedström *et al.* 2008: 714).

On the other hand, Boudon (1993) and Cherkaoui (2000 and 2008) interpretations point to the role of *meso* and *micro*-level mechanisms in Durkheim's work (Boudon, 1993). In this line, Durkheim's theoretical claims would also clearly include the concept of "emergent effects" i.e., a non-additive micro–macro transition where—due to the interdependence of the micro-level units—the whole is more than the sum of the parts (Sawyer, 2005: ch. 6; Macy and Willer, 2002: 1–3)[2]. At the micro-level, Boudon (1993) recognizes action-based explanations in "Suicide". He points out that, according to Durkheim, suicide rates are

higher during upward trends in the economic cycle, given that anomie is also higher. In terms of generative mechanisms, Boudon goes on, this means that in a (perceived) stable phase, individuals expect future gains alike to those experienced today, while in a positively (perceived) stage they expect higher gains. During periods of growth therefore, a higher number of people believe they can be economically successful and thus become risk-prone. Due to a classical congestion mechanism however, competition increases more than available "prizes" and the number of "losers" rises accordingly. Thus, Boudon creates an abstract model that shows the micro generative processes of the macro-to-macro relation between the trends of the economic cycle and the suicide rates.[3]

Cherkaoui (2008), on the other hand, supports the idea that Durkheim in the "Rules of Sociological Method" underlines an holistic explanation, unsuitable to the mechanism-based approach, while in "Suicide" he applies a meso-level approach orientated towards the explanation of emergent macro-level phenomena (2005: 48). But is perhaps in the "Elementary Forms of Religious Life" that the idea of "emergent effect" finds its best illustration:

> But it is in the Formes élémentaires that he makes his thinking clear in the fullest manner by focusing on the emergence of religion in ancient society (Cherkaoui 2008, 35).

Indeed in the "Elementary Forms of Religious Life", the "sacred" as a normative phenomena is stripped down to a series of social factors which qualify a particular model of *social interaction* known as "ritual" (Collins 2004; Goffman 1999). In a ritual interaction actors are concentrated on a "common focus of attention" in face-to-face situations. This generates an emotional crescendo, which becomes shared cultural meaning reinforced by subsequent interactions. By producing a situational collective effervescence, then, group belonging is created, along with enthusiasm, individual energy, symbolic objects that represent and reinforce *ingroup/outgroup* identity as well as a sense of justice, moral standards and legitimate sanctions in the case of violation of these standards.

As aptly underlined by Cherkaoui (2008, 33–39), in the "Elementary Forms of Religious Life" Durkheim considers non-additive processes

only in the micro–macro link: religion and morality are emergent *macro*-level phenomena generated through ritual interactions. On the other hand, the macro–micro link is based on atomistic and additive mechanisms, which refer to the socialization process in an *over-socialized* manner. It should be possible to argue that Durkheim considers the role of meso-level explanations also in the macro–micro link, with an emphasis on corporations, associations and "intermediate bodies". Unfortunately the way Durkheim dealt with this issue remained mainly philosophical (Maddox, 2003) and—with the due respect—with little explanatory power.

Therefore, Durkheim's meso and micro level mechanisms work asymmetrically in the micro–macro and macro–micro processes. In trying to overcome this "theoretical asymmetry", we will show that interaction rituals are also key in macro–micro *immergent* processes (see footnote 1). This will help to apply analytical sociology methodological framework to the genesis of social identities, comparison criteria among actions and recognition processes.

2. WHAT DO MACRO–MICRO MECHANISMS EXPLAIN?

In analytical sociology, macro–micro mechanisms mainly aims to explain the genesis of the elementary components of the action's schema: desires (what agents want to do), beliefs (what agents think they can do), opportunities (what agents can actually do). In the simplest case the relevant factors are the constraints and information *exogenously* available in a situation of "free exchange" (Abell, 2003). To add complexity to this baseline model, analytical sociology *endogenizes*—primarily through social interactions and social networks—the genesis of the action's schema components (beliefs, desires and opportunities, see Hedström, 2005). So far, so good. But what is left out from this conception of macro–micro mechanisms?

In a dialogue with Peter Blau, Coleman (1993: 63; see Udehn, 2001, 302–304) in the context of a discussion around his *opus magnum* "Foundations of Social Theory" (1990) admits: "Were I writing the book over again, I would give considerably more attention to

this macro–micro relation". In a work pointedly entitled "Analytical Sociology: Bringing Culture and Identity Back In", it is argued that in analytical sociology:

> Too little attention is being paid to the macro-to-micro link. In Hedström (2005) this link is almost exclusively represented in the form of structures constituted of interacting individuals, and we argue that more than dyadic interactions must be taken into account in order to fully see the importance of the macro level on individual-level actions. (…) We argue for the importance of macro-level entities such as culture, social categories, and groups—all of which have so far been largely omitted from analytical sociology. Secondly, we criticize the persistent focus on intentionality as the driving force of social action (Edling and Rydgren 2014) .

Edling and Rydgren's criticism, in our opinion, hits the nail on the head. To be clear, it needs to be underlined that—unlike Edling and Rydgren—more complex forms than dyadic interactions are taken into account in analytical sociology and include both interaction with a "social whole" (e.g., the average behaviour of the group) as well as social networks topology (Hedström and Bearman, 2009: 14).[4] Nonetheless, the *mainstream* conception of macro–micro mechanisms does not pay enough attention to the issue of social identity, recognition processes and comparability among action (Pizzorno, 2006). The bottom left-hand corner of the Coleman Boat often implies *pre-formed agents*, without accounting for the cultural and symbolic dimensions of social processes that *constitute* agents as such.[5] For example, often agents reflexively ask themselves about their own preferences: they do not merely desire something, but rather they question how much these desires are "correct", as in the case of a "life-plan" (Frankfurt, 1971). These meta-preferences refer back to those social processes that preside over the genesis of trans-individual evaluation criteria or categories,[6] which in turn refer back to the origin and function of comparability schemes between actions and agents.[7]

As Edling and Rydgren (*ibidem*) suggest, these issues force us to step outside the intentionalist paradigm, acknowledging:

> the growing awareness that the human mind works in two different modes. The first is automatic, quick, reflexive and largely unconscious, while the other is controlled, slow, reflective and conscious.

To give credit to these processes, Edling and Rydgren refers mainly to social and cognitive psychology, while alike arguments derive from behavioural economics (Manzo, 2013). Our claim is that the model of ritual action offers better analytical tools, particularly in order to underline the interdependence between structural elements and cultural processes. Before outlining rituals as mechanisms, however, we shall dig deeper at the premises and consequences of the theory of recognition and identity, to show how these issues force to overcome the primacy of intentionality in the explanation of meaning.

The sociological theory of recognition and identity is built on a fundamental distinction between theories that are intentionalist and outcome-oriented on the one hand and those that are contextualist and reception-oriented on the other. As Pizzorno states:

> The intentionalist type considers that the explanation of actions is reason oriented. Propositional attitudes, i.e. propositions defining beliefs and desires of the agent, are deemed to establish the cause of the action. Rationality is predicated on intentionality. (…) Intentions are therefore considered to be internal states of mind, which occur prior to expressing themselves in behavior. The agent must always be aware of his intention and, if asked, would claim them as his own. (…) On the other side, we have theories of social action which define the meaning of an act not through reference to the intentions of the agent, but through the process by which this act comes to be understood by the participants in the social context where it takes place. In other words, the meaning of an action is what those to whom the action is directed at understand it to be. (…) Rationality does not bear on the intention of an agent, but rather on the reception of the action among the participants (Pizzorno, 2006: 374).

The concept of recognition, Pizzorno goes on, fills two lacuna of intentionalism. The first lacuna is the absence of a vertical/temporal dimension: action is always embedded in time and agents need therefore to face "value uncertainty". At time $t+1$ the *recognition circle* which guarantees the meaning and hence the value of any particular choice may be different to the one at time *t*. The robustness of the vertical/temporal dimension depends on the horizontal dimension of ties that constitute the recognition circle (*ibidem*: 390): the action does not conclude with the agent's satisfaction, but depends on the *judgment* that others (the recognition circle) gives to the action's results. It follows, as Pizzorno underlines, that the "elementary unit of social action" is not the decision-making process at the individual level, but rather the *interaction* process with others who judge our actions. For this reason, the very possibility of social action depends on contextual conditions, such as the stability of the circle of recognition.

These social conditions are at the basis of the *persistence* of individual goals through time. Judgment refers to the ways people name or evaluate one's action, and the meaning of acting is what those to whom the action is directed at understand it to be: "one undertakes to see the world as others do—not because the benefit of doing so outweighs the cost, but because that is *the way of being in the world* with these people" (Loury, 2002: 44, emphasis added). When the sources of naming change, semantic and ontological uncertainty occur (Lane and Maxfield, 2005) and agents face a sort of "symbolic tsunami" (Pizzorno, 2006: 392). Here, the threat derives from the lack of meaning and of ontological stability that can be attributed to a course of action and its potential outcomes. For example, the social change created by the French revolution had its founding moment in the "Taking of the Bastille", 14th July 1789 (Sewell 1996). But what everyone now defines as the "Taking of the Bastille" was actually the result of a complex pattern of interactions, which evolved from declarations, debates and symbolic behaviour. This went along with changes both in the behaviour of individuals and in the stand of the various groups and parties involved. The sociological theory of recognition thus urges to introduce those mechanisms according to which: "people make relations in the short run, while relations make people in the long run" (Padgett and Powell, 2012, 3; see also Abbott 2001)[8]. In the

following we will show how the extension of the Durkheimian theory of rituals helps to outline these mechanisms.

3. INTERACTION RITUALS AS MECHANISMS

As we anticipated in section one the Durkheimian theory of rituals has potentially a double value. First of all, as Cherkaoui noted, this theory holds the keys to find micro–macro effects and therefore the tools to look at the micro foundations of *emergent macro phenomena*. Let us consider in this regard the reformulation of "a ritual theory of class culture" made by Collins in his early works (1975, 1988). This theory unable us to define the micro–macro mechanisms that explain the emergence of a bi-dimensional social stratification system. The ritual-like interactions within a social class consolidate normative orientations, moral standards, in-group or out-group boundaries and agent's behaviour accordingly. These interactions, we could argue, consolidate a pattern of recognition rules, establishing mutual judgments and reducing uncertainty. Thanks to these rules others can be identified and honoured as "persons" within the circle of recognition: they become "one of us" and therefore worthy of trust and respect. In this way, recognition rules become part of the identity of the person and generate a sense of belonging to a group. At the same time, these rules may also become a source of inequality. This process is alike to that which allows us to recognize—by its diagonal movement on a chessboard—a piece of carved wood that is more or less like a bishop. What would happen if this piece of wood moved a little diagonally and then a little circled? By breaking the rules it would no longer be recognized as a chess piece. In the same way, individuals would slip outside a particular recognition circle, or would not enter it, if they cannot act according to its recognition rules. The inequality created by this exclusion is at variance from vertical inequality dictated by power. While the latter distinguishes individuals, for example, by whether they give or take orders (Coleman, 1990), the exclusion from rituals differentiates individuals on whether they have the "resources" to belong to the group or to the same collective identity. For example, this might be because they are part of the same culture, experience the same feelings, talk the same language, share the same life-style and so on and so forth. According to

Collins (*ibidem*), it is precisely this second dimension that produces the crystallization of status groups, which along with power relationships generates at the macro-level a bi-dimensional social stratification pattern.

Secondly, the theory of rituals allows to single out *immergentist macro–micro mechanisms*, thus overcoming both an under-socialized and over-socialized explanation of values (Granovetter, 1985 and 2000). As such the theory of rituals has a wider scope than that used by Durkheim, overcoming the asymmetry between macro–micro and micro–macro explanations of values that—as Cherkaoui pointed out—denotes Durkheim's theoretical approach. To grasp this point we have to illustrate in some detail the use of networks in a contextualist perspective, showing the link between the theory of networks and that of rituals.[9]

4. STRUCTURAL EQUIVALENCE AND COMPARABILITY AMONG ACTIONS

The use of networks in a contextualist perspective assumes that individual action takes place in a situation where participants understand certain acts as being reasonable, predictable and meaningful in the light of shared beliefs: that is within a recognition circle. As we have assumed, without the agent's relational positioning "value uncertainty" occurs: this does not depend on the agent itself, but rather concerns the judgment that the outcome of his choice will receive from other people. What are my aims? What are they worth? The value of the choice is predicated upon the judgment of people other than the agent.

Here is it the close connection with the theory of rituals. A ritual interaction is precisely what allows us to be recognized as persons: it is a representation worth of symbolic value, as it shows that we are part of a *common belonging*.[10] This representation is not an "intentional action directed towards an aim", but rather "an expression of being and belonging" whose symbolic meaning it's acknowledged by other agents for its fitting to an institutionalized system of rules, regardless the intentions of the agent.

Now, as we know, a representation as a symbolic expression needs an appropriate context: a cut in a canvas is not always recognized as a

work of art! As Searle (2005) wrote, X stands for Y *in Z*, because X to be a symbol for Y needs a context Z. In our argument this context is given by a particular configuration of networks, those which define the circle of recognition. When the specific relational configurations of a recognition circle emerge from the complex inter-play between individual and collective responses to different symbols, interactions acquire the value of a ritual: i.e., they become representations of states of *being* and *belonging*. These network configurations make up the ritual situation.

There is therefore a two-stage process which can occur at different speeds. In the first stage a complex interplay of micro behaviours generates, in part even stochastically, a particular network configuration that defines a circle of recognition. This is an emergent micro–macro effect that may occur independently from individual intentions. In the second stage, within the circle of recognition, agent's interaction acquires the intrinsic value of a ritual interaction, signalling the shared belonging to the circle. This is an *immergent* macro–micro mechanism that works as well independently from individual intentions. When and how a network structure constitutes a circle of recognition?

On the basis of Harrison White's analysis (1992; 2001; Azarian, 2005 and 2010), we could argue, a ritual situation emerges from networks with structurally equivalent nodes. White too—in a contextualist perspective—states that an individual outside specific relational conditions will be prey to "value uncertainty". Positioned at the intersection of multiple networks that link different parts of the social space, the agent is bombarded erratically by different and contradictory inputs. His position within a network is so specific that "cultural rules" are too general and therefore useless as a guide to action. A support could however be the comparison with other peers, who are in the same structural situation and exposed to the same flow of inputs and perceived constraints. By comparison, actors who feel they are structurally equivalent can work out what "normality" is and/or which the standards for evaluating actions are (Azarian, 2005). It is by this comparison and control that they can generate and reproduce their own (organizational) schemes, which White calls "disciplines". These schemes become a stable mode of action, support a collective identity and allocate social roles accordingly. Within a given discipline, agents improve their control over erratic flows

and uncertainty that afflict them: they share a relatively stable system of constraints and norms, and acquire a stable framework for action accordingly: a discipline therefore allows behaviour's comparability. Disciplines shape individual identity anchoring it to a wider identity, i.e., to a collective profile or to an order of worth (Stark, 2009). Whoever fits the measures and standards dictated by such a scheme is recognized as a member of the discipline: through the effort to act according to a given discipline, agents *acquire comparability* (Azarian, 2005).[11]

In White's argument the achieving comparability, even if he is using a different terminology to that customary in ritual theory, is the core of ritual action: namely being included in a social circle. Stepping outside White's language, we can say that the schemes of a discipline are a group of rules that—when they are applied—become the symbol of collective identity to which agents can refer to stabilize their own individual identity. In this way, the scheme's rules are not only technical, but rather they are constitutive rules: moving diagonally in a certain context a piece of wood can be recognized as a bishop and it can be *compared* to other chess pieces. Being subject to the rules of a discipline allows us to become comparable, which can be equated to a ritual that displays to others that we are part of a collective identity. White's theory allows to place the meaning of ritual action into the concept of achieving comparability, shedding light on the structural conditions of such action.

Here we can appreciate the social conditions behind the emergence of a ritual representation, which we previously defined as the first stage of a ritual action. White explains how comparability is an emergent possibility of the network. It is generated by the presence of groups of nodes (agents), who have similar ties with other nodes. Groups of nodes that are structurally equivalent, make up the *discipline* that make comparison and ritual representation possible. The first stage of ritual action is therefore made up by the emergence of these groups from the interaction of individuals responding to an almost infinite variety of stimuli, in a way that is both technically rational but also emotionally sound. When interactions stabilize in blocks of structurally equivalent nodes, then, agents recognize their reciprocal similarity and "comparison" are made accordingly.

White, by clarifying the structural conditions of comparison, shows the *immergent* nature of ritual action: that is the social conditions for reciprocal recognition, ritual action, collective and individual identity. Being recognized as subjected to a set of common rules and so acquiring comparability are capacities that the agent acquires *non-intentionally* thanks to his "immersion" in particular structural conditions. These conditions are alike to being able to classify colours when there is light. We cannot state that someone who sees colours *intends* to see them. Rather, seeing colours is a *way of being* inherently human, which occurs under appropriate light conditions quite independently of the agent's intentions! All in all, for White, ritual action and therefore collective and individual identity, come out from particular configurations of social networks. This occurs when—contingently on a chaotic variety of stimuli—the flow of inter-personal links forms blocks of structurally equivalent nodes.

5. THE CONSTITUENTS OF SOCIAL STRUCTURE

An important implication of the link between the theory of networks and that of rituals concerns the constituents of social structure. We saw in section three how, according to Collins, ritualized interaction makes social structure more complex by forming in-groups and out-groups boundaries. To the vertical inequality between classes based on power, therefore, it needs to be added the horizontal inequality of status that refers to the possession (or not) of qualities with an intrinsic value to a group.

White's discipline also introduces inclusion and exclusion criteria on the basis of a relational category, i.e., a "block". An individual *i* exists and is defined as being similar to *j* as each have similar relationships with third parties. Hence each discipline constitutes itself into a node of a larger structure which can be identified by block modelling techniques (Azarian, 2005: 93). The ties between nodes made up by the discipline define the role structure on which stands the overall social organization. This structure can be considered as a catnet (White, 2008; Hedström, 2005: 118), that is a network of social categories or roles which are stable even when individuals change. Therefore an agent is part of a given

category or role through his *relations with others* and not thanks to his own characteristics. If a "colleague" A is substituted with B, B becomes the new "colleague". At the same time, the participants of each discipline, exactly because they are comparable, can compete to distinguish themselves. The organizational schemes that characterize each discipline thus give valuation ordering or rankings, on which a hierarchy is generated.

In this way, White's theory states that it would be difficult for a vertical class dimension to exist without producing horizontal segmentation of in- and out-groups and vice-versa. Therefore, following White's argument, it is unlikely that social space will be mono-dimensional[12]. The vertical differences of class, as we have previously illustrated, depend on the acquisition of resources (money, power) which condition the ability to command others. White holds that the achievement of ability to command is difficult unless it is within the context of organizational schemes in a discipline—i.e. within blocks of agents who are structurally equivalent and able to measure this equivalence. For example, the accumulation of money by the sale of products are only possible in marginal cases, where quality comes from the information carried by the price. Usually however, acquisition and competition need particular stable relational structures which define quality standards (Karpik, 2010) and in which group of peers monitor each other on the basis of a shared criteria of quality (Boltansky and Thevenot 2006).

Besides, consolidating a relational structure that is stable enough for agents to be able to compare to each other constitutes the macro context for ritual action. It follows that, apart from marginal cases, action aimed at acquiring money through the production of goods is carried out in a relational context that gives this exchange a symbolic value of representation that is shared by the collective identity. It is unlikely that this action is purely a technical-instrumental accomplishment (Granovetter 2000). In short, if the metaphor of the perfect market is kept within its structural scope conditions, then we can say that rarely will *homo oeconomicus* control uncertainty in order to be intentionally successful, at least without also becoming (unintentionally) *homo sociologicus* whose actions have *per sé*—for their fit to a set of institutional rules and not for their intentional consequences—a symbolic worth.[13]

A symptom of the difficulty of separating *homo oeconomicus* from *homo sociologicus*, i.e., the intrinsic bi-dimensional nature of social structures, becomes paramount during a crisis when structures "unravel" (e.g., due to globalization when a mixture of financial and economic reality is difficult to single out). This reaches the point in which agents can no longer answer the question "What business am I in?", so they can no longer compare themselves through shared quality criteria. In these cases White suggests that agents, rather than discard their evaluation criteria (e.g. worlds of quality or orders of worth) which is part of their own role and identity, will "manoeuvre" their ties using coupling–decoupling techniques to find a new niche in which their evaluation criteria still works (Barbera and Audifredi, 2012). Therefore, the capability of being recognized as subjected to a set of common rules and so "being comparable" may be achieved thanks to a non-intentional macro–micro "immergent" mechanism. Nonetheless, if and when structures unravel an intentional effort to preserve this capability is needed and when this effort of achieving comparability fails or becomes too demanding a new niche is needed and intentional adaption may require complete reconversion to other "orders of worth" (Bellanca 2007).

DISCUSSION

To conclude we will discuss to two additional consequences, relating to the implications of our arguments about macro–micro immergent mechanisms for the conceptualization of social change, namely: a) crisis and transformation of citizenship's models; b) processes of social innovation.

a) Rituals and Models of Citizenship

As Hedström illustrated (2005) a catnet like those generated by White's "discipline" can influence contagion processes.[14] This was modelled within a catnet in which each category can, even if very modestly, influence (or be influenced by) one or more others category. Here is the apparent paradox according to which particularism can give rise to universalism: "[i]n a social environment in which agents are divided socially and only more or less pay heed to what their peers do, large

and widespread organizations are more likely to arise than in similar but undifferentiated social environments" (Hedström, *ibidem*: 121). This mechanism allow us to model how network blocks can aid or impede shared, and indeed, widespread conceptions and patterns of citizenship in modern societies. We can therefore reduce the scope of the strongly culturalist explanations of these models, which invoke *ideologies* of equalitarianism, individualism and meritocracy, work ethics, "familism", and prejudices based on various forms of race, gender, class or ethnic discrimination.

To dig deeper into this issue, we can look at the concept of social citizenship that have emerged in western industrialized societies over the thirty years, since the second world war. As we know, these societies used to have large markets operating in Fordist business sectors centred on durable consumer goods. They were generally stable and relatively simple societies.[15] From White, we could say that the hierarchical structure of large firms have, along with public bureaucracies, created—for a large amount of the population—a stable catnet made up of a small number of equivalent nodes of class positions, with clear and differentiated power.

On the basis of Hedström's diffusion mechanisms we can then argue that this social stratification has favoured—from the categories laid out by mass production—a "spontaneous" spread of homogenous evaluation criteria. According to these criteria, the quality of goods and consumption is defined by the standards of "democratic luxury" that gave to all workers, even with medium–low incomes, the possibility to conform to the life style and correspondent life course that originally characterized only the urbanized middle class. These standards of mass consumption have then configured a series of "disciplines" which are incorporated in the nodes of equivalent positions or roles. For example, in some European countries these disciplines have placed the men/fathers—whether civil servants, self-employed, professionals, managers or blue-collars—as the "male bread-winner" and wife/mothers as the main "care-giver". The sons and daughters "normally" were part of the student category until adolescence, or even later.

By taking on the "masks" of family roles, thus disappearing behind these roles, women and men of different social classes have been able to

show their belongingness to the same collective identity. Respecting family roles have therefore become a ritual representation, worth of intrinsic value (Collins, 1988: 244). By following the timing and rhythm that regulate the access to these roles, i.e. transition to adulthood, individuals were able to recognize themselves as "moral persons". In this way, the discipline of mass production has supported the post-war model of citizenship of the western societies. Securing the conditions (both material and immaterial, individual and collective) of access to this discipline would allow agents to be considered as citizens. If and when these conditions fail to materialize, the penalty would be the loss of a sense of common belonging, social cohesion and reproductive capacity of society in itself. This analysis sheds light on the troubles of countries such as Italy, in which the Fordist model of citizenship has progressively broken down without being replaced by another general model.[16]

b) Ritual Action and Social Innovation

These issues are even clearer when considering the implications of a wider purpose of ritual theory to the analysis of social innovation (for an introduction to the topic, see Moulaert *et al.,* 2013). In the fifth section, we saw that except for marginal cases, the money-goods exchange occurs in structured contexts that give these actions a symbolic value. We concluded that buying a good is rarely simply a technical act, but often has a ritual intrinsic value. This holds true as well for power, prestige, influence and those resources that, even more than money, are not free to work impersonally as media of systemic integration, without involving the life-world of agents. For instance, power is symbiotically bond to the power holder (Habermas, 1984: ch. 7) and the same applies for other generalized symbolic media of exchange such as prestige and influence (La Valle, 2001). Therefore, empirical distribution of any individual's achievement as well as the conditions and rules to obtain it (e.g., the number and type of "positions" in an organization, salaries level and career's selection criteria), have their own intrinsic value.

From the "discipline" which regulates the acquisition processes, a "grammar" of the results and their possibility conditions emerges. This grammar is applied in the precisely articulated daily activity of "naming", which has a wide range of different labels. As a consequence, a

change in the distribution of results and/or in their possibility conditions, even if Pareto-superior, can change the recognition rules, thus generating frustration and uncertainty. This is true not only for changes due to globalization, new technologies, urbanization, immigration, economic crises and social mobility. It can also be due to policies that are intentionally aimed at improving greater equality. Policies which change a consolidated distribution of results and their acquisition conditions can broke the web of symbols which agents rely on in everyday ritual representation

Here we have a paradox: redistributive action—necessary to extend the chances of recognition—also questions cultural frameworks from which recognition itself is derived. This dilemma seems un-resolvable unless a new consensus around shared values is created. If the fabric of values is being unravelled by public policies, then this very action needs to generate the new structural conditions of ritual interaction that will remedy the problem it has created. In light of this, Appadurai (2001), for example, reflected on the relevance of the role of collective protest by the poor in India. To underline the relevance of collective action, Appadurai referred to those interaction components that give them a ritual character. In this perspective he—referring to extremely poor contexts in urban India—pays close attention to describe the poor's organizational strategies. These may also be initiated by third parties, which translate requests to intervene on a local social scale, often for elementary needs, into appropriate collective action for self-mobilization and self-management by the poor. Appadurai shows these events may be capable of generating fields of informal interaction, in which a symbolic repertoire is generated, transforming immediate needs and their responses. These responses to simple stimuli (hunger, cold, illness) take an intrinsic value for the participants to collective action. For the rest of the world they stand for the aspiration of "being together with others in a certain way".

To conclude, the crisis of the Fordist model of social citizenship as well as the processes of social innovation in India show how the explanation of social change *implies* to overcome the asymmetries of Durkheim's theoretical approach. In particular, it urges to consider both the emergent micro–macro mechanisms of social phenomena, as well as the macro–micro immergent mechanisms that enable

mutual recognition and the construction of individual and collective identities.

NOTES

1. The concept of "immergence" considers processes of macro-micro downward causation (Cambell 1974; Andrighetto *et al.* 2007; Conte *et al.* 2007). Individual properties are non-intentional consequences of the context in which agents are immerged: to anticipate a metaphor we will use later, if the light is switched on I see colors around me independently of my intentions. This is in line with the Latin roots of *im-merge*, that means to disappear by entering into any medium, such as the moon into the shadow of the sun. The concept of immergence we will use in this paper therefore differs from the so-called "second-order" emergence or "incorporation" in which the emergent effect is represented in the producing system, and this representation contributes to replicate the effect (Conte *et al.* 2007, 42).
2. Similarly, Bearman defends Durkheim's stance on the structure of social networks, which contributes to explain the various suicide rate between, for example, Catholics and Protestants (Bearman, 1991).
3. Boudon also proposes another interpretation of Durkheim's findings, through the lenses of a cognitivist model of rationality (1993). In the previous example, the model used to explain the macro-regularity is both utilitarian and rational. An important feature of this model is "consequentialism", i.e., actors choose according to the consequences they expect from their actions. But this is not a general model of action. Coming back to "Suicide", Durkheim states that suicide rates are greater in the second part of the economic cycle than in the first. Boudon's interpretation is that this is due to the fact that each actor hold her own beliefs according to the "position" at any given point in time. Given that the functional form of economic cycle is not linear, positional reasons would therefore lead to over or under-optimistic beliefs, generating greater or lesser delusion accordingly.
4. In other words, the so-called "endogenous social effect" includes much more than just "dyadic interaction". On this issue, see also: Abell (2003) and Mansky (1993).
5. This does not mean that the issues of social identity and recognition processes are excluded from the frame of analytical sociology. For instance Manzo (2012, 44), argues that individual's desire to possess and to defend a well-defined social identity has to be considered as a "reason triggers" mechanism, based on cognitive "good reasons" à la Boudon (*ibidem*, 52). Nonetheless, actors—and their identity-based wishes—are always pre-formed entities (e.g. actors *have* an identity and social circles just *confirm* or *disconfirm* it): "Pizzorno's idea—Manzo argues—of "recognition circles" implies (that) the actor needs others with similar views in order to be reassured about his identity, which is what makes belief and reason formation possible. (…) (I)t is postulated that actors need to search for potential sources of (more or less) local social consensus" (*ibidem* 45; emphasis added).

Therefore, the emphasis is for the consequences of those actions triggered by the wish the preserve one's identity. This schema fits well with Harrison White' idea of "manoeuvring", which we will refer to in the paper. This idea underlines the intentional coupling and decoupling of ties in order to preserve or adjust the "order of worth" or "discipline" in face of a crisis. On the contrary, the aim of this paper is to endogenize the genesis of the contexts in which actors, regardless their intentions, may or may not identify; where their interaction, despite their will, can't avoid to be structured as a ritual; where the human need of being recognized emerges and develops over and above individual intentions. In our opinion, the two theoretical frameworks and the resulting research programs can be fruitfully integrated.

6. This is a classical topic in Durkheim's thought: "Finally, there is a seventh kind of category that I will call "Durkheimian categories". These are the fundamental concepts that facilitate the normal functioning of human society. For Durkheim, they included space, time, causality, substance, genus, quantity, and *personhood*, among others" (emphasis added) (Schmaus 2006, 430). This topic leads to the classical Hobbes–Locke–Hume debate about the distinction between "freedom of will" and "freedom of action" which states that even if we are free to act, we do not freely choose our wishes and wants (Sciolla 2008, 101).
7. Reflexivity may concern both the means or the ends of one's action. We will not elaborate this distinction here.
8. This statement looks very much like Coleman's statement about the dynamic relation between levels: actors contribute to create the macro-level on the short run, and, once it is created, the macro-level retroacts on actors, and so on. We are grateful to the editor of this book for this point.
9. The adoption of a contextualist perspective instead of an intentionalist one has deep implications for the use and conception itself of social networks (Azarian 2010). For instance, within an intentionalist framework, a variety of different reasons can make an individual join other similar individuals (Hedström 2005, 118), thus creating groups and segmenting social networks by categories. There is a wide spectrum of reasons for these homophile tendencies: from creating an oligopoly, to decreasing the costs of being close to dissimilar individuals (Schelling 1978) or to the possibility of exploiting positive externalities by living among similar people (Bellanca, 2007). Others well-known mechanisms refer to rational imitation connected to isomorphic dynamics (Powell and DiMaggio, 1991; Fligstein, 2001), path-dependence and evolutionary processes (Nelson and Winter, 1982). In all these cases agents are always *pre-formed* entities able to anticipate the value of their actions and outcomes on the basis of their preferences and beliefs. On the contrary, in a contextualist perspective ties constitute agents as such: outside these ties and groups agents can't figure out the value of their choices.
10. We can also note that the basic elements of Collin's rituals (physical co-presence, exclusion of outsiders, participants' common focus of attention, synchronicity of shared activity, collective effervescence) can be deduced from the representational feature of these kinds of interaction. Indeed, representation is like a dance and so, as Pizzorno points out, gestures are "complete in themselves" and not

the means to reach an aim. The dance needs partners, and the focus of attention is on movement, which creates emotional energy. Furthermore the moves of the dance have a rhythm because "they relate to other presences", which "need to be synchronized" (Pizzorno, 2008).

11. According to White, a discipline could be generated and reproduced also by means of an active agency among participants who have their niche of structurally equivalent nodes made by means of coupling–decoupling techniques. Besides, we argue, disciplines can emerge—in part even stochastically—through a complex interplay of micro-behaviours.
12. Our argument leads to consider the overlapping areas among the three main disciplines theorized by White: interfaces, arena and council. But we cannot dig deeper into this point here.
13. The roles thus created may be understood as "masks" (Pizzorno, 2008). These masks allow any human being who wears them (being subjected to the discipline of the various blocks) to hide his more peculiar individual characteristics which would make him unique and therefore different: so the agent gains comparability and is recognized as a person (person's original Latin meaning is exactly that: a mask).
14. Hedström (2005, 118) came to this conclusion by simulating the contagion model build by Äberg (2000), based on Schelling (1971 and 1978)
15. As aptly noted by Roberta Sassatelli in reading a draft of this paper, this interpretation of a "simple" and well-separated society may be biased towards the more "central" kinds of people of the Fordist phase (such as "male bread-winner") or specific social categories (such as "social class"), while leaving out "marginal" subjects (women) and key social dynamics ("social movements"). This is indeed, at least in this first example. But we are also considering the collective behavior of poor and marginal people in India, showing that the Catnet-Ritual-Recognition triplet fits well also with the less culturally hegemonic categories of whatever society.
16. In these countries well-known phenomena such as women and youth unemployment, familism, the restrictiveness of the housing market, the downgrading of social mobility and regional disparities (to mention only a few), have variously interfered with the transition to adulthood (Schizzerotto, 2002). In these societies, transition to adulthood is supported by private and family-related forms of capital (Barbieri, 1998), making the creation of ritual and public interactions much more difficult than before (Negri and Filandri, 2010).

REFERENCES

Abbott, A. (2001) *Time Matters*, Chicago, University of Chicago Press.

Abell, P. (2003) "On the prospects for a unified social science: Economics and sociology", *Socio-economic Review* 1(1): 1–26.

Åberg, Y. (2000) "Individual social action and macro level dynamics: A formal theoretical model", *Acta Sociologica*, 43: 193–205.

Andrighetto, G. Campennì, M. Conte, R., Paolucci, M. (2007), "On the immergence of norms: A normative agent architecture", in *Proceedings of AAAI Symposium, Social and Organizational Aspects of Intelligence*, Washington DC.

Appadurai, A. (2001) "Deed democracy: Urban governmentality and the horizon of politics", *Environment&Urbanization*, 13(2): 23–44.

Azarian, R.G. (2005) *The General Sociology of Harrison C. White. Chaos and Order in Networks*. London: Palgrave MacMillan.

Azarian, R.G. (2010) "Social ties: Elements of a substantive conceptualization", *Acta Sociologica*, 53: 323–338.

Barbera, F. and Audifredi, S. (2012), "In pursuit of quality. The institutional change of wine production market in Piedmont", *SociologiaRuralis*, 52, 3: 311–331.

Barbieri, P. (1998) "Regolazione istituzionale e redistribuzione dello stigma. Stato, mercato e reti sociali nei processi di avviamento al lavoro come fattore di esclusione sociale", *Rassegna Italiana di Sociologia*, 2: 249–72.

Bearman, P. (1991) "The social structure of suicide", *Sociological Forum*, 6(3): 501–524.

Bellanca, N. (2007) *L'economia del noi*. Milan: Bocconi.

Boltanski, L. and Thévenot, L. (2006) *On Justification. Economies of Worth*. Princeton, NJ: Princeton University Press.

Boudon, R. (1993), "Toward a synthetic theory of rationality", *International Studied in the Philosophy of Science*, 7, 1: 5–19.

——— (1998), "Social mechanisms without black boxes", in P. Hedstrom and R. Swedberg (eds), *Social Mechanisms*. Cambridge: Cambridge University Press: 172–203.

Campbell, D. (1974) "Downward causation in hierarchially organised biological systems", in F. Ayala, T. Dobzhansky, (eds), *Studies in the Philosophy of Biology*. London: MacMillan.

Cherkaoui, M. (2000) "La stratégie de mécanismes générateurs comme logique de l'explication", in J. Bacheler (ed.), *L'acteur et ses raisons. Mélanges pour Raymond Boudon*. Paris: Puf, pp. 131–151.

——— (2005) *Invisible Codes: Essays on Generative Mechanisms*. Oxford: The Bardwell Press.

——— (2008), *Durkheim and the Puzzle of Social Complexity*. Oxford: The Bardwell Press.

Coleman, J. S. (1987) *Microfoundations and Macrosocial Behaviour*, in Alexander, J. C. *et al.* (eds), *The Micro–Macro Link*. Berkeley, CA: University of California Press: 153–173.

——— (1990), *Foundations of Social Theory*. Cambridge, MA: The Belknap Press of the Harward University Press.

——— (1993),"Reply to Blau, Tuomela, Diekmann and Baurmann", *Analyse und Kritik*, 15: 62–69.

Collins, R. (1975), *Conflict Sociology. Toward an Explanatory Science*, New York, Academic Press.

——— (1988), *Theoretical Sociology*. Orlando, FL: Harcourt Brace.

——— (2004), *Interaction Ritual Chain*. Princeton: Princeton University Press.

Conte, R. Andrighetto, G., Campennì, M., Paolucci, M. (2007) "Emergent and immergent effects in complex Social systems", in *Proceedings of the AAAI* ,07.

Edling, C. and Rydgren, J. (2014) "Analytical sociology: Bringing culture and identity back in", *Sociologica*, Forthcoming.

Esser, H. (1993), "The rationality of everyday behaviour: A rational choice reconstruction of the theory of action by Alfred Schutz", *Rationality and Society*, 1(5): 7–31.

Fligstein, N. (2001), *The Architecture of Markets*. Princeton: Princeton University Press.

Frankfurt, H. (1971), "Freedom of the will and the concept of reason", *Journal of Philosophy*, 68: 5–20.

Gambetta, D. (2009), "Signalling", in Hedström, P. and Bearman, P. (eds), *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press, pp. 168–194.

Geertz, C. (2000) *The Interpretation of Cultures*. New York: Basic Books.

Goffman, E. (1999) *The Presentation of Self in Everyday Life*. Peter Smith Publisher Incorporated.

Goldthorpe, J. H. (2000), *On Sociology*. Oxford: Oxford University Press.

Granovetter, M. (1985), "Economic action and social structure: The problem of embeddedness", *American Journal of Sociology*, 91(3): 481–510.

——— (2000), "A theoretical agenda for economic sociology", in M. Guillen, R. Collins, P. England & Marshall Meyer, (eds.), *The New Economic Sociology: Developments in an Emerging Field*. New York: Russell Sage Foundation, pp. 35–59.

Gross, N. (2009). "A pragmatist theory of social mechanisms", *American Sociological Review*, 74: 358–379.

Habermas, J. (1984) *The Theory of Communicative Action*, vol. 2. Boston: Beacon Press.

Hedström, P. (1998) *Rational Imitation*, in P. Hedström and R. Swedberg (eds), *Social Mechanisms*. Cambridge: Cambridge University Press: 306–327.

——— (2005) *Dissecting the Social: on the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.

Hedström, P., Liu, K., Nordvik M.K. (2008), "Interaction domains and suicide: A population-based panel study of suicides in Stockholm, 1991–1999", *Social Forces*, 87, 2: 713–740

Hedström, P. and Udehn. L. (2009) "Analytical sociology and the theorie of middle-range", in Hedström, P. and Bearman, P. (eds), *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press, pp. 25–50.

Karpik, L. (2010) *Valuing the Unique: the Economics of Singularities*. Princeton, NJ: Princeton University Press.

La Valle, D. (2001) *La ragione dei sentimenti. Una teoria dello scambio sociale*. Rome: Carocci.

Lane, D. and R. Maxfield (2005) "Ontological uncertainty and innovation", *Journal of Evolutionary Economics*, 15(1): 3–50.

Lindenberg, S. (2000), *The Extension of Rationality: Framing versus Cognitive Rationality*, in *L'Acteur et ses Raisons. Mélanges en l'honneur de Raymond Boudon*, in J. Baechler *et al* (eds.), Paris, PUF: 168–204.

Loury, G. (2002) *The Anatomy of Racial Inequality*. Harvard: Harvard University Press.

Macy, M. W. and Willer, R. (2002) "From factors to actors: Computational sociology and agent-based modelling", *Annual Review of Sociology*, 28: 143–66.

Maddox, M. (2003), "Durkheim beyond liberalism: Liberal accommodations of group rights", in Paul Morris and Helen Greatrex (eds), *Human Rights Research*. Wellington: Victoria Human Rights Programme, pp. 71–102.

Manski, C. F. (1993) "identification of endogenous social effects: The reflection problem", *The Review of Economic Studies*, 60(3): 531–542.

Manzo, G. (2010) "Analytical sociology and its critics", *European Journal of Sociology*, 51(1): 129–170.

——— (2012) "Reason-based explanations and analytical sociology. A rejoinder to Boudon", *Revue Européenne des Sciences Sociales*, 50(2): 36–65.

——— (2013), "Is rational choice theory *still* a rational choice of theory?", *Social Science Information*, 52(3): 361–382.

Moualert, F., MacCallum, D. Mehmood, A. Hamdouch, A. (2013), *The International Handbook On Social Innovation, Collective Action, Social Learning and Transdisciplinary Research*, Cheltenham, Edward Elgar.

Nelson, R. and Winter, S. G. [1982], *An Evolutionary theory of Economic Change*, Cambridge Mass., Harvard University Press.

Negri N. and Filandri M. (eds) (2010), *Restare di ceto medio*, Bologna, Il Mulino.

Padgett, J.F. and Powell, W.W. (eds.) (2012), *The Emergence of Organization and Markets*, Princeton, Princeton University Press.

Pizzorno, A. (2006), "Rational choice and collective action", in S. Turner and M. Risjord (eds.) *Handbook of the Philosophy of Science*, Vol. 15: Philosophy of Anthropology and Sociology: 373–398.

——— (2008), *Sulla maschera*, Bologna, Il Mulino.

Powell, W.W. and DiMaggio, J. (eds.), (1991), *The New Institutionalism in Organizational Analysis*, Chicago, Chicago University Press.

Rydgren, J. (2009), "Beliefs", in Hedström P. and Bearman P. (eds.), *The Oxford Handbook of Analytical Sociology*, Oxford, Oxford University Press: 73–93.

Sawyer, K. (2005), *Social Emergence: Societies As Complex Systems*, Cambridge, Cambridge University Press.

Schelling, T. (1971), "Dynamic Models of Segregation", *Journal of Mathematical Sociology,* 1: 143–186.

——— (1978), *Micromotives and Macrobehaviour*, New York, W. W. Norton & Company.

Schizzerotto, A. (ed.), (2002), *Vite Ineguali. Disuguaglianze e corsi di vita nell'Italia contemporanea*. Bologna: Il Mulino.

Schmaus, W. (2006) "Categories and classifications in the social sciences", in S. Turner and M. Risjord (eds), *Handbook of the Philosophy of Science*, vol. 15: Philosophy of Anthropology and Sociology: 429–458.

Sciolla, L. (2008) "La forza dei valori", *Rassegna Italiana di Sociologia*, 1: 89–116.

Searle, J.R. (2005) "What is an institution", *Journal of Institutional Economics*, 1(1): 1–22.

Sewell, W.H. (1996) "Historical events as transformations of structures: Inventing revolution at the Bastille", *Theory and Society*, 25(6): 841–881

Stark, D. (2009) *The Sense of Dissonance. Accounts of Worth in Economic Life*, Princeton, NJ: Princeton University Press.

Udehn, L. (2001), *Methodological Individualism*. London: Routledge.

White, H.C. (1992), *Identity and Control. A Structural Theory of Social Action*. Princeton, NJ: Princeton University Press.
——— (2001), *Markets from Networks. Socioeconomic Models of Production*. Princeton, NJ: Princeton University Press.
——— (2008), "Notes on the constituents of social structure *Soc. Rel.* 10—Spring '65", *Sociologica*, 1, doi: 10.2383/26576.

13

THE FOUR RULES OF "GOOD SOCIOLOGY"

Raymond Boudon
Académie des Sciences Morales et Politiques, Paris

In 1990, James Coleman, one of the leading American sociologists of his generation, published a thick important book, *Foundations of Social Theory*. As the very title indicates, the ambition of the book was to give sociology a solid ground. While Coleman recognized the importance of the works of classical sociologists, notably Weber and Durkheim, he presented his work as very critical toward them. In a nutshell, he proposed to make the so-called *rational choice theory* the basic paradigm able to make sociology a genuinely scientific discipline. He even tried to translate this paradigm into institutions, as the prestigious journal *Rationality and Society* or the seminar led by Gary Becker and I had the privilege to attend at Chicago University in 1986. Becker received the Nobel Prize for economics thanks to his application of the rational choice paradigm, the traditional pillar of economics, to sociological problems as marriage or discrimination.

Many sociologists perceived Coleman's endeavour as a reaction against the lack of identity and solidity of sociology that appeared as increasing since the 1960s. In the 1980s, two diametrically opposed orientations dominated sociology: structuralism and constructivism. The former saw social actors as conditioned by social and cultural forces,

the latter as led by more or less arbitrary mental constructions. The former appeared as inspired by scientism rather than by the scientific *ethos*, the latter by a dogmatic version of scepticism, if I may use this oxymoron. This situation explains possibly the audience Lepenies' (1985) theses enjoyed: that science is neither a science nor an art; that classical sociologists as Weber and Durkheim were naïve when they claimed that sociology could be a science like any other.

A more recent reaction against the high diversity of contemporary sociology appeared in the last years under the label *analytical sociology* (Manzo, 2010). A comment on the back of the collective book edited by Pierre Demeulenaere under the title *Analytical sociology and social mechanisms* claims that analytical sociology is simply "good sociology" (Demeulenaere, 2010). This raises the question: what is good sociology? Is it grounded on ultimate principles? If so, which ones?

Mohamed Checkout's writings on Durkheim and Weber, as well as his criticism of Coleman's work suggest an answer to these questions. They show that Durkheim and Weber proposed more or less implicitly a paradigm able to make sociology a science like any other and that this paradigm should be distinguished as well from the *rational choice paradigm* as from the *naturalistic paradigms* according to which human behaviour would be essentially the effect of dispositions resulting from social, cultural, psychological, biological, psychobiological or neuronal forces. As Cherkaoui suggests through his brilliant analyses of a host of empirical analyses from Durkheim and Weber, the paradigm includes four rules. Unfortunately, handbooks are seldom aware of the presence of this paradigm, not only in Durkheim's and Weber's works, but in the work of contemporary sociologists. All handbooks agree more or less that Durkheim and Weber are the fathers of sociology, while they curiously tend to insist on the divergence rather than the convergence of their conceptions of sociology. By so doing, they increase the confusion around the identity of sociology.

METHODOLOGICAL SINGULARISM

The first rule claims that sociology can reach the goal of any science, i.e. explaining phenomena appearing at first view as opaque, under

the condition that they are well defined. Émile Durkheim raises in his *Suicide* a number of questions dealing all with *singular* social phenomena: Why are the suicide rates of women lower than the suicide rates of men? Why are the suicide rates of Protestants higher than the suicide rates of Catholics? Why are the suicide rates of bachelors higher than the suicide rates of married people? Once he had succeeded explaining these and other singular phenomena about rates of suicide, Durkheim attempted to synthesize his explanations in the broader theoretical framework of his three –actually four- types of suicide.

Explanations of a host of intriguing phenomena characterize in the same way Max Weber's work. Thus, he wonders in his *Old Judaism* why Pharisians believed in the immortality of the soul, while Sadducees did not (Weber, 1920). In his essay on the protestant sects in America (Weber 1920), he wonders why Americans are much more religious than Englishmen, Germans or Frenchmen and discovered a cause that earlier analysts of the American religious exceptionalism, as Adam Smith (1776) or Tocqueville (1840), had missed. The accumulation of the numerous explanations of puzzling singular religious phenomena he gathers in his *Essays in the Sociology of Religions* (Weber, 1920) proposes on the whole, as Durkheim's *Elementary Forms of Religious Life* (Durkheim, 1912), a new way of explaining religious phenomena.

The Austrian economist Ludwig von Mises has coined the notion "methodological singularism" to christen the rule according to which a scientific explanation can bear exclusively on singular phenomena (von Mises, 1949). He wanted to distinguish by so doing between *explanation* and *interpretation*. Thus, a complex sequence of events as, say, the 1789 French Revolution, can be *interpreted* rather than properly *explained*, because it is by essence ill-defined. To begin with, nobody knows when it ended, nor when it got started. To many historians, it ended with Bonaparte taking the power in 1800. To François Furet, it ended in 1870 with the Parisian Commune. Before von Mises, Georg Simmel had also made the point in a great book that some topics can give birth to *interpretations*, others to *explanations* (Simmel, 1892). Thus, there can be no unique final biography of historical major actors, because a biography rests always upon a selection and differential weighting of a huge number of facts, the selection and weighting depending necessarily upon the

sensitiveness of the historian and of his time. By contrast, there can be a unique explanation of a singular phenomenon as to why women commit less frequently suicide than men or why Americans are more religious than Europeans. In a word, to von Mises, methodological singularism is a necessary condition of any genuinely scientific explanation.

This rule applies, not only to sociology, but also to all social sciences. That an Austrian economist coined the notion of *methodological singularism* is easy to understand. At the end of the 19th century, German and Austrian economists give the impression of living in distinct intellectual worlds. To the German historical school and notably its leading figure Gustav Schmoller, the main objective of economics was to describe the history of economic institutions. The members of the school had little consideration for the young Austrian marginalist economists who cared rather to explain singular economic phenomena. The conflict has given birth to a famous quarrel, a *Methodenstreit*, which remains famous in the history of economic ideas.

On the side of sociology, the same intellectual quarrel characterizes the history of the Frankfurt school, as an anecdote drawn from Rolf Wiggershaus' monograph on the school illustrates (Wiggershaus, 1995). In the 1930s, several members of the Frankfurt school leave Germany and try to get settled in the American academic world. Max Horkheimer, a leading figure of the school, proposes to develop an ambitious social *Theory*, which he insists to describe with a capital T, a practice usual in German but strange in English, so that the American academic authorities are reluctant to accept his project. A young Austrian sociologist by the name of Paul Lazarsfeld helped him efficiently. But as soon as the two had gained the recognition of the American academic world, they could not help expressing the deep intellectual distance they felt between one another. The main reason of this distance is that Lazarsfeld endorsed methodological singularism, while Horkheimer Grand theory rejected it.

By accepting implicitly the principle that von Mises called later the principle of methodological singularism, the founding fathers of sociology followed a principle governing all scientific disciplines. Thus, in his current work, a biologist deals normally with singular phenomena, as when he explores the behaviour of some virus. He feels certainly

concerned with broader questions, as the origin or essence of life, but they do not directly belong to his research goals.

Max Weber follows the same principle as, say, biologists when he explains for instance why Pharisians believed in soul immortality while Sadducees did not. This comes from the fact, he contends, that Pharisians were mainly shopkeepers and businessmen. To them, the notion of the equity of exchanges was a professionally crucial value. Consequently, the idea that the actions of people on this earth would be equitably rewarded after their physical death was palatable to them. Sadducees were by contrast the category from which the political and religious elites were drawn. Sadducee priests followed the Bible teaching: the main aim of people is to have a long life, while respecting the religious rules is the means to this goal. The idea of a life after death was foreign to them. They tended to see the immortality of the soul as a strange idea imported presumably from India or Persia.[1]

Before Weber, Tocqueville (1840) had explained the Indian belief in reincarnation in the same way: as a symbolic expression of the idea that in the long run, at the end of the cycle of reincarnations, good and bad behaviour would be equitably rewarded.

Many sociologists do not follow the methodological singularism principle. Thus, the German sociologist Ulrich Beck (Beck, 1992) does not care to explain singular puzzling social facts, but is rather concerned with characterizing the main features of modern societies: they would be risky societies (*Risikogesellschaften*). In the same way, the English-polish sociologist Zygmunt Bauman describes modern societies as "liquid": as societies that would have lost any moral and intellectual fixed point (Bauman, 2008). Earlier, Gustave Le Bon reached an audience that Weber or Durkheim never reached, by describing the societies of his time as characterized by the reign of "crowds" (Le Bon, 1895).

Singular phenomena can take several forms, notably: of regularities of the type *if A, then B* (as: Catholics commit less frequently suicide, Americans are more religious than Europeans, to reciprocate a gift too hastily is not recommended in some circumstances), of trends (as: in the Western world penal sanctions have become progressively less heavy, people tend to give overtime less attention to religion), of single well-defined events, etc.

METHODOLOGICAL INDIVIDUALISM

So, *methodological singularism* is the first rule common to all scientific explanations in the social sciences. Max Weber and Joseph Schumpeter, who mentions in his *History of Economic Analysis* that he has worked a while for Weber, identified a second rule. They called it *methodological individualism*. Popper and Hayek then popularized the notion. Max Weber saw methodological individualism as the basic principle of what he called "comprehensive sociology": "comprehensive sociology, in the way I take this expression, considers the isolated individual and his action as its ultimate unity, as its 'atom'"[2] (Weber, 1922b). Weber's text indicates that methodological individualism is the basic principle of comprehensive sociology, because individual actions are the only possible causes of social phenomena. Unfortunately, handbooks often interpret wrongly methodological individualism as ignoring that social actors are embedded in institutions and social contexts or as including the utilitarian principle according to which interests would lead all human actions.

It should be noted that the notion of *comprehensive sociology* itself was also and is still often misunderstood. Weber makes this clear through the careful parenthesis ("in the way I take this expression") he introduces in the previous quotation. This shows that he was aware that many social scientists gave in his time the notion of *comprehensive sociology* a holistic rather than individualistic meaning. But the confusion between the individualistic and the holistic meaning is far from being clearly perceived today. Thus, some handbooks characterize as "comprehensive" the type of *holistic* sociological theories that attempts at describing, say, the features characteristic of given societies or eras, as Dilthey and Burckhardt in the past or Ulrich Beck and Zygmunt Bauman nowadays.

On the whole, a history of the misunderstandings around the notions of "understanding" (*Verstehen*) and "methodological individualism" waits still to be written! A famous article entitled the "operation called *Verstehen*" has permanently attracted a great deal of attention (Abel, 1948). Its thesis is that sociology has nothing to gain from the operation called *Verstehen*. It derives from Abel's implicit postulate that reasons cannot be the causes of behaviour, or alternatively that sociologists are doomed to explain behaviour with the help of a-rational factors:

of psychological, biological or social forces that would determine the representations and behaviour of men in the back of their mind. The thesis rests in other words on the debatable naturalistic theory of knowledge according to which a scientific explanation should always mobilize *material* causes. This is certainly true of the natural sciences, but not of the human sciences (Bronner & Sauvayre, 2011; Boudon 2012: Ch. 7).

Against these misunderstandings, Weber defended clearly the idea that explaining singular social phenomena was the natural objective of sociology and that, as social phenomena can have no other causes but human individual actions, explaining them implies finding out the causes of these individual actions, more precisely identifying the reasons actors have to do what they do or to believe what they believe. Weber explains why Pharisians believed in soul immortality and why Sadducees did not by trying to disentangle their reasons. The causes of this macroscopic difference lay in the reasons that led ideal-typical Pharisians to find the idea palatable and ideal-typical Sadducees to see it as strange. This does not imply that the real individuals would have clearly perceived the reasons that caused their attitude toward the issue. Such reasons are generally *metaconscious* rather than *conscious*, because actors concentrate normally on the goals and means of their actions rather than on their reasons. But the latter are not *unconscious*, in the sense that actors can more or less easily retrieve them.

Against the view defended by most handbooks, Durkheim endorsed implicitly also the principle of methodological individualism, notwithstanding his rather obscure statements on the relations between the individual and social levels, in modern vocabulary: on the micro-macro link (Borlandi, 2011). His statements on the subject have by far not the clarity of Weber's. But, while his theoretical developments are obscure and possibly inspired by his goal to impose sociology as a hard science excluding subjective causes, his empirical analyses lean upon a conception of the micro-macro link similar to Weber's, as Cherkaoui (1990; 2011) indicates. Thus, in his *Suicide*, Durkheim explains the macroscopic fact that Protestant commit more frequently suicide than Catholics by introducing the assumption that ideal-typical Catholics are led by moral principles imposed by the Roman Church, while Protestant have to devise by themselves the moral principles they should follow

by deciphering the messages of the Holy Scriptures under their own responsibility, so that they are more likely exposed to doubt and anxiety when existential problems confront them and hence more likely to find in suicide a solution to their problems.

Or, to take an example drawn from Durkheim's last book on the *Elementary forms of religious life*, he explains the fact that "soul" is a universal notion that moreover appears as much more resilient than other religious notions, by showing that the notion has a clear symbolic meaning for any individual in any society. In any society, he claims, people have the strong feeling that some types of behaviour are legitimate or not, good or evil. In our modern vocabulary, any individual has a sense of *values*. Moreover, he has the feeling that the values he endorses are a basic component of his self, but that he has not devised them. Consequently, he cannot avoid experiencing a strong feeling of the duality of his self. To Durkheim, the notion of the *soul* should be interpreted as a symbolic expression of this duality. Contemporary sociologists of religion, as Thomas Luckmann (1967), confirm this explanation. I have personally shown that modern empirical research confirm Durkheim's thesis. The *World values survey* (Inglehart, 1998) shows that, while people become in most countries more and more sceptical about the dogmas and concepts of religions, "soul" appears as a highly resilient notion (Boudon, 2012: Ch. 4; 2013).

The notion of soul is in other words a metaphor identifying a complex reality. Generally, people understand the metaphors in use among them because they identify more or less clearly the complex reality underlying them. As Cherkaoui (1998) has stressed, Durkheim makes the illuminating point that metaphors or proverbs are more frequent in traditional than in modern societies because in traditional societies people are embedded in a common cognitive context, while in modern societies the division of labour, the increasing distinction between roles and other factors have the effect that people live in highly diverse cognitive contexts. That *Mutter Courage* speaks by proverbs in Brecht's masterpiece expresses her high popularity among the soldiers. As Hardin (1997) has shown, using a special vocabulary or syntax or special metaphors is a means for the members of a sectarian group to stress their collective identity and their distance from the surrounding society.

I could also have taken examples from Durkheim's *Division of Labor* to illustrate his implicit use of methodological individualism. The secular trend that we observe toward a substitution of civil for penal judiciary sanctions and toward increasingly lenient sanctions is caused by the fact that ideal-typical citizens tend to prefer a less heavy sanction as soon as it appears as equally efficient. Unfortunately, handbooks retain only from Durkheim's analysis that the increase in social "density" and in the division of labour facilitates this process.

Cherkaoui (1998; 2011) presents another example illustrating eloquently Durkheim's individualistic methodology. Other things equal, divorced people commit less often suicide than widows or bachelors when divorce rates are higher because, living in societies where divorce is likely, married people cannot see marriage as a serious protection against the existential difficulties they meet, so that they are less affected than widows when their couple is destroyed.

Methodological individualism is the social scientific version of a general rule valid for all sciences. When a biologist, say, observes a correlation between the regular consumption of some food and the frequency of some disease, he will normally look for the elementary mechanisms responsible for the correlation. Methodological individualism recommends in the same way to explore the elementary individual microscopic causes responsible for the macroscopic social phenomena. Paradoxically, these elementary individual microscopic causes are more easily retrieved in the case of the social than of the natural sciences. In the former case, they take the form of reasons parameterized by the context ideal-typical actors are embedded in. In the latter case, researchers are rarely sure to have reached the ultimate elementary level. In Weber's terms, the "atoms" of the social sciences are ultimate components, while the "atoms" for instance of physics have components of still lower level.

BEYOND THE INSTRUMENTAL CONCEPTION OF RATIONALITY

The notion of "comprehension" (*Verstehen*) was limited in Weber's mind to the explanation of individual actions. It encapsulates the third essential principle underlying the vision Weber and also Durkheim notably

had of sociology as science. To them, the *causes* explaining that any ideal-typical individual behaves the way he does or believes what he believes lie exclusively in the *reasons* as to why the individual behaves the way he does or believes what he believes. More precisely, an ideal-typical individual endorses some belief as soon as he has the impression that it is grounded on a system of reasons arousing in his mind the impression that they are acceptable and that he fails to see an alternative system that would lead to a different belief. And he undertakes some action if he has the feeling it is an adequate means to fulfil some goal he has reasons to follow. It should also be added that these reasons are in the general case parameterized by the context: The Sadducees had not the same reasons as the Pharisians to believe in the immortality of the soul.

Durkheim is even more explicit than Weber on this point. One of the basic postulates of his explanation of religious beliefs is that people believe what they believe because they have strong reasons to believe it. He is very critical against the sociologists of his time who, as Max Müller or Lévy-Bruhl, analyze religious or magical beliefs as illusions generated in the mind of people by occult social or psychological forces. He is very ironical against the idea that beliefs can be explained as illusions or, as he rather says, "hallucinations", indicating by this word that he is very critical of sociologists who see people's minds as led by obscure natural forces. When a young child plays with his puppet, the observer has easily the impression that the child believes that his puppet is alive. But the child would be surprised if the puppet would suddenly bite him. Even a cat, adds Durkheim ironically, understands very quickly that a ball is not a mouse. He plays with it a while, but leaves it as soon as he understands that it is not a mouse (Durkheim, 1912). Durkheim would certainly have rejected the postulate abundantly declined in the Marxist tradition that consciousness would be basically "false", i.e. that social actors would be normally blind on the reasons as to why they do what they do or believe what they believe. This central postulate of *structural sociology* led it to a dead-end. As Jean-Daniel Reynaud (2002) has claimed, the objective of sociology would have been according to Pierre Bourdieu to "deliver to human beings the meaning of their actions" (*Redonner aux hommes le sens de leurs actes*).[3] To Weber and Durkheim

by contrast, actors are led by reasons, though they are only half conscious of them simply because they normally concentrate on their goals rather than their motivations. This trivial observation is far from the postulate that people would systematically endorse false representations of their motivations and reasons under the effect of the social structures. Hence the task of the analyst is to disentangle the reasons actors *have in mind*. Durkheim and Weber endorse this postulate in all their analyses.

A simple example from Weber illustrates the idea that beliefs should rather be explained by the reasons that, given the context, individuals have to endorse them. Thus, an ideal-typical Western observer is surprised by the fact that "primitive" people seem to believe that their rain rituals are able to produce rain, but that they rub pieces of wood against one another to produce fire does not intrigue him. The context explains the difference: a typical modern Westerner is aware that kinetic energy can be transformed into thermic energy. For this reason, the primitive fire maker does not surprise him, while he sees the behaviour of the rainmaker as strange. As to the "primitive" himself, he has no reasons, claims Weber, to make this difference: he sees rainmaking and fire making as practices equally grounded in theories in which he believes but that ideal-typical Westerners spontaneously see as "magical" (Weber, 1922b).

In this example, Weber introduces implicitly a powerful theory of rationality. This theory breaks with the current view on rationality by seeing rationality as *cognitive* and not only *instrumental*. The example conveys the crucial idea that explaining an individual act, belief or behaviour means determining the reasons that have inspired the act, belief or behaviour in the mind of the individual, given that reconstructing these reasons implies that the observer should be aware of the relevant features of the social and cognitive context in which the individual is embedded.

Cognitive rationality is also at work in the example from Durkheim that Cherkaoui evokes on the correlation between rates of divorce and of suicide: in a society where divorce is frequent, married people cannot be confident that the cooperation and protection they derive from marriage will last. Being *aware* of this fact, they do not *see* marriage as a reliable protection against existential problems, so that their level of

their anxiety and hence their propensity to suicide in case of divorce increase to a lesser extent than the level of anxiety of widows.

I have personally proposed to formalize Weber's theory of rationality and called it the *Theory of ordinary rationality*, in order to distinguish it from the so-called *Rational choice theory*, the theory of rationality most currently accepted today in the social sciences (Boudon, 2012; 2013). The difference between the two is that according to *Rational choice theory* the reasons moving social actors would be basically egoistic and instrumental, while the *Theory of ordinary rationality* claims that they can be egoistic but also impersonal, and instrumental, but also cognitive. Thus, the reasons explaining that Westerners perceive the behaviour of primitive rainmakers as strange but not the behaviour of fire makers are cognitive and not instrumental, impersonal and in no way egoistic.

As it claims a competence limited to the explanation of the means people use to reach their goals, *Rational choice theory* suffers from a major shortcoming: it is doomed to explain the objectives, values and beliefs of people by postulating the existence of conjectural causes operating in the back of their minds. Rational choice theorists call generally these conjectural causes *frames* or *frameworks*. But the same explanatory scheme underlies many other notions as *mentality* (as with Lévy-Bruhl's *primitive mentality*), *habitus* (in the particular use of this traditional notion by structuralism), etc. Some social scientists treat these causes as mere data, while others propose to derive them from psychological, social or biological in many cases highly conjectural mechanisms. For this reason, Popper (1976) has convincingly argued that, as long as the social sciences will insist using such conjectural notions, they will necessarily stand on clay feet.

A clear distinction should be made in this respect between dispositional concepts, as *primitive mentality* in Lévy-Bruhl's or *habitus* in the sense of structuralist sociology, which provide *ad hoc* pseudo explanations of behaviour, and the parameters describing the context of behaviour. That the "primitive" do not know the laws of the transformation of energy is a *parameter* distinguishing their context from the observer's context. That early Christians did not make a distinction between natural and supernatural phenomena is a parameter distinguishing their context from the new context science created when it

introduced the idea all phenomena are governed by laws. Such *parameters* are empirical data, while *dispositional variables* are intrinsically *ad hoc* constructs. Coleman's (1990) famous boat fails to make this crucial distinction, since it collapses the two categories into the single category of "individual attitudes/dispositions".

Dispositions do of course exist, but one should distinguish physical and mental dispositions. The former are difficult to change. Thus, unlearning to walk or to ride a bicycle is practically impossible. The latter are exposed to changes in the context.[4] Moreover, the effect on behaviour of personality features, a subcategory of *habitus*, is highly unpredictable. As Simmel (1892) has shown, an authoritarian education can produce a liberal or an authoritarian personality and vice versa. Empirical studies confirm that the relations between personality features and behaviour are weak and variable from one study to the other.[5] Tocqueville, Weber, Durkheim or Simmel all display a sound agnosticism toward dispositional variables. To them, the context defines the parameters of individual actions and beliefs: it makes that individuals have reasons to do or believe this or that. They never propose an a-rational explanation of beliefs and actions, whereby they would be the mechanical effect of occult social, biological or psychological forces. This explains for one part the solidity of their work. In fact, dispositional variables were included into the sociologist's toolbox under the strong and lasting influence of behaviourism, an intellectual move that had the pretention to make psychology a hard science.

The Theory of ordinary rationality I propose to derive from Weber's work rests finally upon a basic principle I propose to call the *cognitive equilibrium principle*. It says that people believe that X is true, acceptable, good, legitimate, etc. as soon as they have the feeling that X rests upon a set of acceptable reasons and do not see a set of reasons that would make Y more acceptable.

Beside *methodological singularism* and *methodological individualism*, the *cognitive equilibrium principle* is the third principle I see as defining implicitly the paradigm used by the founding fathers as well as by many modern sociologists, a paradigm defining actually the core of contemporary *analytical sociology* and offering a better foundation to social theory than Rational choice theory.

The three principles are of course used, not only by Durkheim and Weber, but also by many classical and modern sociologists. Simmel is for instance often considered as practicing a kind of sociology very different from Durkheim and Weber. The phenomena he is interested in are beyond doubt different. His style is different. But he uses the same principles. Why for instance do we feel as shocking that, when somebody has received some gift, he gives the impression he is impatient to reciprocate (Simmel, 1900)? We feel so, because giving has the function of expressing a feeling of closeness, friendship, recognition, etc. toward the Other. Now, if the honoured person reciprocates too hastily, he transforms the gift into an exchange. This familiar norm can be explained by cognitive, not by instrumental rationality. In the circumstances where gifts are simultaneously exchanged, other subtle norms prevail, as not to hurt the Other by offering a gift of much higher value than he could himself give. In still other circumstances, as in the case of the potlatch explored by Marcel Mauss, the exchange of gifts is a means to rank socially the participants. On the whole, explaining the various norms applied to gifts implies disentangling the cognitive reasons on which they rest.

A contemporary social scientist, Bruno Frey (1997), has shown that people accept more easily to have nuclear waste buried on the territory of their county when they are *not* offered a compensation for the risk involved than when they are. Considerations of interest and instrumental rationality cannot explain this puzzling empirical finding, drawn from a survey conducted both in Switzerland and Germany, while cognitive rationality can: if a compensation is proposed to me, I perceive the offer as a bad bargain. If compensation is not offered, I perceive the offer as a proposal to help the national community. Obviously, since it hurts interests, a sacrifice can be more easily explained with the help of cognitive rather than instrumental rationality.

Hilton Root (1994) uses the same three rules when he explains why in the eighteenth century street protests against the price of bread were rare in London and frequent in Paris, although it was more favourable to consumers in Paris. The reason for the difference is that the political power was concentrated in the hands of civil servants in Paris and that Parisians *knew* it, while the members of the British parliament derived

their power from the provincial landowners who had elected them, so that the people in London *knew* that they had little chance to influence them. I have myself shown that street protests are traditionally much more frequent in France than, say, in Britain or Germany, because of the much more highly concentration of the political power at the executive level (Boudon, 2012: Ch. 6; 2013). This explains also that everybody in France understands the expression *le pouvoir de la rue*, which has no real equivalent in German or English.

That the instrumental theory tends to be dominant in the social sciences is illustrated by the huge literature devoted to the question as to why people vote, although no individual vote can make the difference in the final outcome. None of the theories using instrumental rationality has successfully explained the phenomenon, while cognitive rationality easily explains it.

The history of ideas on social theory is fascinating. In the 17th century, La Rochefoucauld (1964 [1678]: 424–425), as later Bentham or James Coleman, develops in his *Maximes* the view that human behaviour should be analyzed as inspired by interests. To him, "all human virtues get lost in interest, as rivers get lost in the sea", "We praise virtue and blame vice only by interest"[6]. In Weberian terms: normative beliefs should be explained by instrumental rationality. But La Rochefoucauld is also aware that social theory should take cognitive rationality into consideration beside interest: "There is an infinity of behaviours that look ridiculous and which rest on very wise and solid hidden reasons"[7]. He identifies clearly the importance of cognitive rationality beside instrumental rationality, as Weber later: "Not interests -material or intellectual- but ideas dominate ultimately human action"[8] (Weber, 1920).

AGGREGATION

A fourth principle, the *aggregation principle*, states that the aggregation of individual actions can produce macroscopic effect that may be unwanted and unanticipated by social actors. In some cases aggregation is simple: everybody in a crowd being in the same situation, all behave in the same way. In other cases, it may be more complex: if students tend to choose medicine because there is presently a shortage of physicians, this

will generate in a few years an overproduction of physicians. By contrast with the three others, this fourth principle is widely accepted and has been widely explored, so that it is useless to say more on it. As Cherkaoui has shown, Weber and Durkheim use it abundantly (Mica, 2012).

SOME FURTHER EXAMPLES

An example from Weber will illustrate the application of the four principle paradigm to the explanation of representational beliefs.

Why, asked Weber, did the monotheistic cults imported from the Middle East, as Mithraism, attracted the Roman civil servants and army officers, while Roman peasants felt deeply hostile to these cults and remained faithful to the traditional polytheistic Roman religion? Their hostility to Christianity was so deep that the word *paganus*—peasant—turned out to be used by Christians to describe the unbelievers: the heathens.

Peasants had trouble accepting monotheism, explains Weber, because the uncertainty characteristic of natural phenomena, which is an essential dimension of their everyday life, did not seem to them compatible with the idea that the order of things could be subjected to a single will: a notion that implies a minimal coherence and predictability. So, the basic reason why the peasants rejected monotheism, beside of course other more or less contingent factors, is that the Roman peasants, as good followers in advance of Popper's falsification theory, had the impression that the monotheistic theory was incompatible with data familiar to them. This analysis explains also that an impressive number of *saints* appeared in the early centuries of our age in the Christian world. Thanks to the saints, Christianity became more palatable to peasants, since they made it a polytheistic religion. Army officers and civil servants had by contrast the feeling that Mithraism and then Christianity expressed symbolically the basic principles on which the political organization of the Roman Empire rested.

The four-rule paradigm can also explain long-term trends. Durkheim (1893) has observed, as we recall, that a secular trend towards more lenient judiciary sentences characterizes Western societies. Also, civil rather than penal law treats an increasing number of types of

misbehaviour. And an increasing number of acts are prosecuted before lower-level courts.

These long-term trends derive mainly from a basic process: when a new type of sentence appears as equally effective in terms of dissuasion as a former one, and also as better from some viewpoint, the ideal-typical individual and hence the public opinion tends to accept it. In other words, a basic two-stage mechanism is at work in this type of long-term processes: 1) production of innovations, 2) rational selection of the innovations. Some factors can facilitate this basic mechanism, as the increasing demographic density and the resulting increasing division of labour, as Durkheim claimed. But the trends can only be explained by taking into account the motivations and reasons of ideal-typical individuals.

The previous analysis can be easily applied to our modern world. Death penalty tends to disappear from modern democratic societies notably because it has been repeatedly shown that its dissuasive power is controversial. It makes judiciary errors irreparable and is obviously cruel. Moreover, the progress of genetic research makes judiciary errors more easily detectable. In the long term these reasons lead ideal-typical individuals to prefer life sentencing to death penalty. Generally, cognitive rationality explains effectively collective moral feelings (Boudon, 2011). As to axiological rationality, it can be held as a variant of cognitive rationality characterized by the fact that at least one statement in the set of reasons explaining that a group holds X as good, bad, legitimate, etc. is of the evaluative type (Boudon, 2012b: Ch. 1; 2013).

Cherkaoui (1998) has stressed the importance of Durkheim's analysis of crowds: why do crowds give the impression that people in the crowd behave as barbarians; why do "harmless burgers become a dreadful monster"[9]? Gustave Le Bon has proposed an a-rational theory of crowd behaviour: when people are inserted in a crowd, their behaviour would be determined by invisible psychological forces that do not operate in normal conditions. Hence Le Bon's idea that crowds would have a psychology of their own, as the title of his book indicates: *Psychologie des foules*. Against this view, Durkheim proposes a rational explanation of crowd behaviour: in a political protest, everybody in the crowd tends to see that he would have to pay a high cost if he would express some

disagreement with the aim the protest follows. Moreover, he joined the crowd because he agrees with the aim and means of the protest. In other words, to Durkheim, the members of the crowd are inspired by cognitive rationality, both representational and normative: *normative*, since they approve the means and ends of the crowd; *representational*, since they clearly see the danger to which they would expose themselves if they expressed some disagreement with the aims and means of the protest. Ettore Scola illustrates beautifully in his sociologically penetrating movie *Una giornata particolare* that people seem to be psychologically different when they cook spaghettis in their kitchen and when they take part in a fascist meeting. But the contrast is an *emergent* effect: people do not follow different psychological rules in their kitchen and in a crowd; they do not become different. The truth is that the kitchen and the crowd parameterize in different ways the context in which individual actors behave. Unfortunately, Le Bon's a-rational theory has always been more popular than Durkheim's rational theory.

The same could be said of the huge literature on magical beliefs. Durkheim has suggested a rational explanation of the beliefs in the efficiency of magical rituals: members of traditional societies believe in their efficiency because they have strong reasons to believe in it. By contrast, Lévy-Bruhl proposed an a-rational explanation of the belief. While Durkheim's theory is more acceptable in the sense that it is made of a set of acceptable statements, Lévy-Bruhl's theory is more commonly accepted (Sanchez, 2007; 2009; Boudon, 2013). Contemporary leading anthropologists as Sahlins (1976), Needham (1972) or even d'Andrade (1989) are closer in this respect to Lévy Bruhl than to Durkheim. Other example: Homans' (1962) explanation of the rules enforced in traditional societies regarding the marriage between parallel and crossed cousins is rational, while Lévi-Strauss's (1949) is a-rational. The latter is more widely accepted, although it is more debatable. A-rational explanations of behavior are more easily accepted because they are more dramatic than rational explanations. And also because they endorse a *naturalistic* view of men, seeing human behavior as determined by material causes[10] (Boudon, 2012: Ch. 7).

One of the important features of the four-rule paradigm and of the cognitive theory of rationality it includes is that it can solve easily many

questions Rational choice theory is unable to solve. I will just mention a few additional examples.

As game theory rests upon the axioms of Rational choice theory, several situations of interaction have no "solution" in the sense that Rational choice theory is unable to recommend any satisfactory line of action. The classical structures of the *prisoner's dilemma*, the *insurance game*, the *chicken game* or the *battle of the sexes* have inspired a huge literature because they have no solution in the frame of Rational choice theory. The ordinary theory of rationality provides by contrast a solution: actors confronted with such structures *understand* that they are trapped in the structure and that the only way to get rid of the trap is to produce or to accept an innovation able to destroy the structure.

The best social thinkers and sociologists have effectively understood that the right "solution" for the actors embedded in such structures is to take steps to the effect of modifying the structure of the game, whenever this is feasible. Thus, Rousseau has well seen in his *Discourse on the origin of inequalities* that the way to solve an *insurance game* is to destroy the structure by introducing legal constraints. Citizens tend to accept them because they *see* that they make them better off. Mancur Olson has shown that collective action can be trapped in an *n-person Prisoner's dilemma* structure, and that the solution is to destroy the structure by some innovation, as the *closed shop* or the production of *selective incentives*. The *crossroads problem* has no solution in the frame of Rational choice theory, since it provides no way of choosing between the two Nash equilibriums the problem generates. The solution is again to introduce an innovation in the form of a priority rule. The interesting side of Axelrod's (1984) book on the repeated prisoner's dilemma game may not be that the *TIT for TAT strategy* provides a solution to the game, but rather that, if actor A cooperates in the first move, this sends to B the *signal* that, if he cooperates, A will likely go on cooperating in the following moves. In other words, Axelrod introduces implicitly the cognitive dimension of rationality beside its instrumental dimension: the players have cognitive reasons to prefer cooperation to defection.

THE SCIENTIFIC TRADITION *VS.* OTHER SOCIOLOGICAL TRADITIONS

In a word, it can be shown that many sociologists from the classical to modern time have followed a paradigm defined by four major principles: *methodological singularism*, *methodological individualism*, the *cognitive equilibrium principle* and the *aggregation principle*. It states that social phenomena are the outcome of individual actions and that the causes of individual actions and beliefs lie in the reasons people have to endorse them. More precisely, people endorse some representational or normative belief if they have the impression that their belief is grounded on a system of reasons that appears to them as better than any alternative system of reasons. As to their actions, they include always beliefs people have reasons to endorse. As the various examples presented in this paper show, the reasons people mobilize are parameterized by the social context. As "primitive" do not know the rules of the transformation of energy, they consider fire and rain making as equally magical.

The four-principle paradigm defines a particular type of sociology. It can be qualified as *scientific* in the sense that it proposes to explain any puzzling social phenomenon by making it the outcome of a set of well-confirmed laws and of factual statements. Now, this is also exactly what any physicist, biologist or astronomer does. In Frey's example, the question as to why people accept more easily to have nuclear waste in their backyard if they are proposed no compensation is solved by using laws as: "people like to be perceived as helpful to their community", "people dislike bargains they perceive as bad", "if a compensation is proposed, people will see the proposal as a bargain", "if no compensation is proposed, people will see the proposal as an opportunity to help the national community". Frey himself did not propose this cognitive explanation, probably because, as he was educated in economics, he had in mind the dominant instrumental conception of rationality. He mentions vaguely the idea that economists should accept that a-rational psychological factors interfere with rationality. But he does not come to the idea that instrumental rationality is a particular form of rationality. Kahneman himself (2011), who is critical of Frey, presents the "biases" revealed by his experiments as a-rational psychological factors.

As Cherkaoui (2005) has shown, the same can be said of Coleman. He misunderstood Weber's theory in his *Protestant ethic*, because of his instrumental conception of rationality. He did not see the importance of Weber's point that Islam had not the same effects on behaviour as Puritanism because Muslims are convinced that they follow God's will as soon as they respect the rituals He prescribed, while Puritans have no other sign of their election by God than the success they meet in their worldly activities.

All sociologists have obviously not followed and do not follow the four-rule paradigm. I have evoked the contrast between the *individualistic* conception of sociology that Weber, Durkheim and many modern sociologists, as the few I have mentioned, illustrate and the *holistic* conception Le Bon, or today by Ulrich Beck or Zygmunt Bauman illustrate. Beside this distinction, several other distinctions have been proposed. There is a *descriptive sociology* that explains that sociology can sometimes be seen as an art rather than a science, as rightly stated by Gary Runciman and Brigitta Nedelmann in an international inquiry on sociology (*Commentaire,* 2011). There is a *critical sociology* according to which social theory should have the main objective to defend some social cause. There is a *theoretical sociology* canonically illustrated by Parsons' *Theory of Social Action* or Lazarsfeld's theory of empirical action: both propose a reflexive approach aiming to identify the criteria that make sociological works more or less convincing. There is a *naturalistic sociology* represented by structural sociology, memetics or sociobiology, or the incantatory sociology that sees its future in the neurosciences. It supposes that human behaviour is the effect of *material* causes, i.e. of social, psychological or biological causes that, as the causes of digestion, would exclude any intervention of the human mind (Boudon, 2012: Ch. 7; 2013). Finally, there is a *words-words-words sociology*, which can be forgotten. On the whole, contemporary sociology gives the impression of being highly diverse. Unfortunately, this leads many handbooks to present sociological theories more or less as opinions and for this reason to juxtapose them in the way grocery products are presented in commercial catalogues.

The four-rule paradigm eliminates the dilemma detrimental to all social sciences according to which they would be doomed to explain

human behaviour with the help either of instrumental or of a-rational factors. It explains the scientific strength of Weber's and Durkheim's work. Checkout's writings invite future handbook writers to revisit the history of sociology, to be aware of the convergence between Durkheim and Weber beyond their points of divergence and also to identify and evaluate the various paradigms sociologists have used over time, leading them or not to a better understanding of social phenomena.

NOTES

1. Albert Assaraf suggests to me that the violence notably of the Roman repression against Jews in 70 with its hosts of innocent victims may have eroded the teaching of Sadducee priests according to which respecting the law would be a promise of a longer life and made more credible the belief in soul immortality (personal communication).
2. "*Die verstehende Soziologie, in unserem Sinne, behandelt das Einzelindividuum und sein Handeln als unterste Einheit, als ihr 'Atom'*".
3. L. Boltanski's (2012) version of false consciousness is inspired by the Marxist tradition: the local embeddedness of people would tend to generate false representations in their mind, while intellectuals would be *free-floating*. But in complex societies people lie in diversely parameterized "beds". This generates more or less complex aggregation effects. Unembeddedness is moreover far from being characteristic of intellectuals. Ordinary people can be unembedded on issues that do not meet their interests: Spectators can be impartial (Boudon 2012, Ch. 6).
4. The notion of *habitus* can be legitimately given an explanatory value provided it can be shown that an actor has reasons to keep or change his *habitus*. Thus, the "primitive" has reasons to keep his *habitus* to consider fire-making as magical. Most sociologists recognize today that the structuralist use of the old notion *habitus* is unacceptable, as Lemieux (2009).
5. Adorno's *Authoritarian personality*, the canonical endeavor to explain behavior by personality features, was severely evaluated already in the time when it was published.
6. "*Toutes les vertus des hommes se perdent dans l'intérêt comme les fleuves se perdent dans la mer*". "*On ne loue la vertu et l'on ne blâme le vice que par intérêt*".
7. "*Il y a une infinité de conduites qui paraissent ridicules et dont les raisons cachées sont très sages et très solides*".
8. "*Interessen -materielle und intellektuelle- nicht: Ideen beherrschen unmittelbar das menschliche Handeln*".
9. "*… des bourgeois inoffensifs deviennent un monstre redoubtable*".
10. Among many others, B. Lahire (2012) illustrates today the lasting propensity of social theorists since Gall, Cabanis or Lombroso and later behaviorists toward naturalism. Leaning on the work of the brain expert J.-P. Changeux, he evokes

a wishful "dispositionalist" sociology, which would be able to detect people's *habitus* thanks to the exploration of their brain. Naturalism has taken various forms over time: thirty years ago structuralism, twenty years ago, sociobiology were presented as the future of sociology, while this role is rather granted today to the neurosciences.

REFERENCES

Abel, T. (1948) "The operation called Verstehen", *American Journal of Sociology,* 54(3): 211–218.

Axelrod, R. (1984) *The Evolution of Cooperation*. New York: Basic Books.

Bauman, Z. (2008) *Does Ethics have a Chance in a World of Consumers?* Cambridge, MA: Harvard Press.

Beattie, J. (1964), *Other Cultures*. London: Cohen & West.

Beck, U. (1992*), Risk Society: Towards a New Modernity*. London: Sage.

Boltanski, L. (2012), *Enigmes et complots*. Paris: Gallimard.

Borlandi, M. (2011) "Durkheim et la psychologie", in Boudon, R. (ed.), *Durkheim fut-il durkheimien?*, Acts of the colloquium organized on November 4 and 5, 2008 by the Académie des Sciences Morales et Politiques. Paris: Colin, pp. 55–80.

Boudon, R. (2011) "The cognitive approach to morality", in Steven Hitlin and Stephen Vaisey (eds), *Handbook of the Sociology of Morality*, Part II. New York/Berlin: Springer, pp. 15–33.

——— (2012), *Croire et savoir: penser le politique, le moral et le religieux*. Paris: Presses Universitaires de France.

——— (2013) *Sociology as Science: An Intellectual Autobiography*. Oxford: The Bardwell Press.

Bronner, G. & Sauvayre, R. (eds) (2011) *Le naturalisme dans les sciences sociales*. Paris: Hermann.

Cherkaoui, M. (2005) *Invisible Codes: Essays on Generative Mechanisms*. Oxford: The Bardwell Press.

——— (1998) *Naissance d'une science sociale*. Paris/Geneva: Droz.

——— (2011) "La complexité du social et la méthode sociologique selon Durkheim," in R. Boudon (ed.), *Durkheim fut-il durkheimien?*, Acts of the colloquium organized on November 4 and 5, 2008 by the Académie des Sciences Morales et Politiques. Paris: Colin, p. 81–98.

Coleman, J. (1990) *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Commentaire (2011) *La sociologie est-elle une science?* Winter 2011: 1001–1094.

D'Andrade, Roy G. (1989) "Culturally based reasoning", in A.R.H. Gellatly, D. Rogers & J.A. Sloboda (eds), *Cognition and Social Worlds*. Oxford: Oxford University Press, pp. 132–143.

Demeulenaere, P. (ed.) (2010) *Analytical Sociology and Social Mechanisms*. Cambridge: Cambridge University Press.

Durkheim, E. (1893) *De la division du travail social*. Paris: Presses Universitaires de France, 1967.
——— (1912) *Les Formes élémentaires de la vie religieuse*. Paris: Presses Universitaires de France, 1979.
——— (1897) *Le Suicide*. Paris: Presses Universitaires de France, 1980.
Frey, B. (1997) *Not Just For the Money. An Economic Theory of Personal Motivation*. Cheltenham: Edward Elgar.
Hardin, R. (1997) *One for All. The Logic of Group Conflict*. Princeton, NJ: Princeton University Press.
Homans, G. (1962) "Marriage, authority and final causes", in *Sentiments and Activities*. Glencoe: The Free Press, pp. 202–256.
Inglehart, R. *et al*. (1998) *Human Values and Beliefs: A Cross-cultural Sourcebook. Political, Religious, Sexual, and Economic Norms in 43 Societies: Findings from the 1990–1993 World Values Survey*. Ann Arbor: The University of Michigan Press.
Kahneman, D. (2011) *Thinking, Fast and Slow*. New York: Farrar, Straus & Giroux.
Katz, E. & Lazarsfeld, P. (2006 [1955]) *Personal Influence*. London: Transaction.
Lahire, B. (2012), *Monde pluriel. Penser l'unité des sciences sociales*. Paris: Seuil.
La Rochefoucauld (1964 [1678]), *Oeuvres completes*. Paris: Gallimard, Pléiade.
Le Bon, G. (1988 [1895]) *Psychologie des foules*. Paris: Presses Universitaires de France.
Lemieux, C. (2009) *Le devoir et la grace*. Paris: Economica.
Lepenies, W. (1985) *Die drei Kulturen*. Munich: Hanser.
Lévi-Strauss, C. (1949) *Les structures élémentaires de la parenté*. Paris: Presses Universitaires de France.
Lévy-Bruhl, L. (1922) *La mentalité primitive*. Paris: Alcan.
Luckmann, T. (1967) *The Invisible Religion: The Problem of Religion in Modern Society*. New York: Macmillan.
Manzo, G. (2010) "Analytical sociology and its critics", *European Journal of Sociology*, 51(1): 129–170.
Mica, A. *et al*. (2012) *Sociology and the Unintended. Robert Merton Revisited*. Frankfurt/Berlin/Berne: Peter Lang.
Mises, L. von (1949) *Human Action*. Yale University Press.
Nedelmann, B. (2011) "Briser la base nationale de la profession", in *Commentaire* (2011), pp. 1053–1054.
Needham, R. (1972) *Belief, Language and Experience*. Oxford: Blackwell.
Popper, K. (1976) "The myth of the framework", in E. Freeman (ed.), *The Abdication of Philosophy. Philosophy and the Public Good*. La Salle, IL: Open Court, pp. 23–48.
Reynaud, J.-D. (2002) "Pierre Bourdieu (1930–2002): restituer aux hommes le sens de leurs actes", *Revue Française de Sociologie*, 43(1): I–IV.
Root, H. (1994) *The Fountain of Privilege: Political Foundations of Markets in Old Regime France and England*. Berkeley, CA: University of California Press.
Runciman, G. (2011) "Précision", in *Commentaire* (2011), p. 1073.
Sahlins, M. (1976) *Culture and Practical Reason*. Chicago: University of Chicago Press.
Sanchez, P. (2007) *La rationalité des croyances magiques*. Paris/Geneva: Droz.
——— (2009) *Les croyances collectives*. Paris: Presses Universitaires de France.

Simmel, G. (1892) *Die Probleme der Geschichtsphilosophie*. Munich: Mohr.
——— (1900) *Philosophie des Geldes*. Berlin: Duncker & Humblot.
Smith, A. (1776) *An Inquiry into the Nature and Causes of the Wealth of Nations*. Chicago, IL: University of Chicago Press, 1977.
Tocqueville, A. de (1840) *La démocratie en Amérique II, in Tocqueville*. Paris: R. Laffont, collection Bouquins, 1986.
Weber, M. (1920) *Gesammelte Aufsätze zur Religionssoziologie*. Munich: Mohr, 1986.
——— (1922a) *Gesammelte Aufsätze zur Wissenschaftslehre*. Tübingen: Mohr, 1965.
——— (1922b) *Wirtschaft und Gesellschaft*. Tübingen: Mohr, 1976.
Wiggershaus, R. (1995) *The Frankfurt School: Its History, Theory and Political Significance*. London: Polity.
Wittgenstein, L. (1967) "Bemerkungen über Frazer's *The Golden Bough*", *Synthese*, 17: 233–253.

14

PURPOSIVE ACTIONS AND THEIR UNINTENDED CONSEQUENCES

Jacques Coenen-Huther
University of Geneva

Some years ago, Mohamed Cherkaoui explored the paradoxical consequences of social action as analysed by Max Weber and Robert Merton (Cherkaoui, 2006). I would like to lean on his theoretical work in order to show that unintended consequences of purposive actions are never the result of the mere juxtaposition of individual behaviours. Even in the elementary case of an individual acting seemingly alone, interactions occurring in his social environment are always involved.

Cherkaoui summarised the main points of Weber's and Merton's theories regarding unanticipated consequences of action in two typologies. The first one, inspired by the works of Max Weber, includes five categories: 1) the inadequacy between means and ends; 2) the structure of interdependence between actors; 3) the inappropriate transfer of a logic of action from one sphere of activity to another; 4) a conflict of rationalities; 5) the emergence of new values in a field of activity (Cherkaoui, 2006: 15–17). Merton's typology is very similar to Weber's one. It includes also five categories: 1) a lack of information leading to a relative ignorance; 2) an erroneous evaluation of the situation; 3) the undue focus

on the immediate consequences of action; 4) a change in the hierarchy of values, provoked by the initial value orientation of the action; 5) a self-destructing prophecy (*ibid*.: 18–20). The mere enumeration of these categories, combined or not, suggests that the consequences of an action are the result—as Merton would put it (1936: 895)—of the interplay between the action, the situation and the conditions in which the action is carried out. These conditions are heavily dependant on the fact that the environment of the action has always systemic properties. This is to say that it is an organized whole made of interdependent parts and showing a specific pattern of relations. A *social* system, however, exhibits characteristics that systems of a lower order do not: it is a *complex adaptative system* (Buckley, 1967: 128) fitted with feedback mechanisms and able to transform itself in the course of a process (*ibid*.: 172–173). The unintended consequences of actions are produced by these mechanisms in which appears the element of "double contingency" involved in any interaction process (Parsons, 1951: 10). Actions and reactions "may cover a considerable range" which includes "possible permutations and combinations" (*ibid*.: 11). Every actor's behaviour—whether individual or collective—sets a problem to which other actors will react according to their own perceptions and possibilities (*ibid*.: 94). Consequently, as a general rule, the more the social context of action is complex and diversified, the greater the opacity of social processes and the unpredictability of their outcome. That opacity exceeds the capacity of cognitive apprehension of individual actors and collective actors alike. That is why the rationality of the actor is always a "bounded rationality", that is a rationality relative to a certain frame of reference (March and Simon, 1958; 1979: 136). Of course, it is the individual actor who can chose the wrong means, miss relevant information, be torn between two rationalities, be subject to a conflict of values, etc. But all this applies as well to decision-makers in charge of collective interests. It means that unintended consequences have to be understood at the individual level in order to be explained as a collective phenomenon. But it means also that structural factors weigh on individual behaviours. All the categories included in both Weber's and Merton's typologies apply to individual as well as to collective actors. At the basic level of interaction between individuals, each actor, according to Parsons, is influenced by two sets

of expectations: role-expectations defining the standards for his behaviour and expectations relative to the probable reaction of the other in the form of positive or negative sanctions (1951: 38). At the level of interaction between collective actors, the range of expectations is similar. The relation between Great-Britain and Germany before World War I illustrates this. In the first years of the 20th century, the Foreign Office anticipated two possible attitudes from Germany. The first one, based on the usual relations between Nation-States of the time, could offer rewards to a policy of moderation. It was the analysis at the basis of the Grey doctrine. The second one, oriented specifically to a rivalry with Britain, ruled out any attempt at conciliation which would be constructed as weakness. This was the position advocated by the Crowe memorandum (Crowe, 1907) which was a source of meditation for Henry Kissinger. Both the individual and institutional expectations are structured by the features of the surrounding social system. Having this in mind, I will describe and analyse in the following pages various examples of actions which led to unintended consequences.

1. THE INDIVIDUAL VERSUS THE SYSTEM

The most striking example of the ubiquitous social constraints on an individual actor is to be found in Hans Fallada's novel, *Jeder stirbt für sich allein* (1965). The story is a good illustration of a case of inadequacy between means and end but also an example of the difficulty to define the end in such a way as to be able to evaluate the means. The actor is the main figure of the novel, the worker Otto Quangel, the social system in which the action is situated is Nazi Germany in wartime, the conditions of the action are those created by a totalitarian system. Otto Quangel, wants to act against the Nazi regime but it is not clear what he wants to accomplish or whether he really wants to accomplish something specifics: he just wants "to do something" For various reasons, collective action is out of the question in his case: in a climate of fears, colleagues, neighbours, even friends or parents cannot be trusted. The only form of resistance he can reasonably contemplate is to write cards drawing the attention of his fellow citizens on the murderous character of Hitler's dictatorship and disseminate them in public spaces. In spite

of his very limited means, his plan is carefully thought: he does not take unnecessary risks and he can expect some real effects of his initiative. He is portrayed as a rational man. Nevertheless, his enterprise proves to be based on false premisses. Quangel assumes that there are widespread discontents among the German population and that the potential for subversion is real. As a consequence, the anonymous messages carried by the cards could have a multiplying effect and help create a new consciousness on the part of those who would read them. Some people would be stimulated to "do something" also, one way or the other. It could be expected that they would start by copying the cards and distribute them further. The sad reality is that these cards, painfully written on a small scale, were hardly read. In the context of terror created by the regime, they did not provoke reflexion but sheer panic. Most of the cards were immediately brought to the police by people whose main concern was not to be suspected of connivance with the author of the messages or of a half-hearted attitude towards the regime. After two years of obstinate efforts, Quangel is finally arrested by the Gestapo. During repeated interrogations, he learns to his dismay that his clandestine action has been totally useless. Out of three hundred cards disseminated in town, less than twenty have not led to an immediate denunciation to the political police. Quangel is then confronted with the complete failure of his solitary action.

The first question to be asked about this scenario is: why does Quangel believe at once that his action will serve a purpose and that the recipients of his cards will react as he hopes they will do. Obviously, he thinks that his countrymen will react as he himself would react in similar circumstances. This means that the expected reaction is for him to be taken for granted. Where does this sense of evidence come from ? It is certainly a lack of information amounting to a relative ignorance in the Mertonian sense: Quangel overestimates the discontents of the population and the will to resist the Nazi regime while he underestimates the consequences of political terror carried out ruthlessly. But no amount of information, adequate or not, can be sufficient to create the motivation to act. The will to be politically active and to resist a murderous system is also the consequence of a certain upbringing and some life experiences. In other words, it is the result of a socialisation

process involving several socialisation agents at various stages in life. An individual's behaviour cannot be explained if the actor is considered as a calculating adult cut off from his formative years. The Gestapo interrogators consider Quangel's anticipations as foolish. For them, such an act of solitary resistance was bound to fail at the onset. We have to deal here with the dilemma of two rationalities. At the beginning, Quangel's behaviour can be analysed within a framework of instrumental rationality. To be sure, the will to resist originated in an emotion—he is mourning the death of his son, enrolled in the Wehrmacht—and not in a cost-benefit analysis. But the expectation was that the action could possibly have the effect that Quangel had contemplated. However, when this expectation proved to be wrong, Quangel's state of mind evolved. He substituted an attitude of axiological rationality to the first one. Then a clash of attitudes became apparent. Facing the unanticipated failure of his action, he finds satisfaction in having acted according to his principles. On the other hand, the Nazi interrogators stuck to an attitude of instrumental rationality: they kept wondering how an individual with so little resources could fancy himself resisting the regime successfully.

In summary, two different elements are involved here: knowledge and values. According to the Weberian and Mertonian taxonomies mentioned by Cherkaoui, there is an inadequacy means-end due to insufficient information and wrong anticipation but also a competition of values. The deficit of information amounts to an erroneous evaluation of the situation. It is the result of a lack of social contacts. But it is not only the fact that Quangel is a rather solitary man; the totalitarian terror inhibits the contacts between neighbours, colleagues, friends and even parents. One has to deal with a disintegrated society in which the only source of information is the State. These structural conditions have an impact on Quangel's choice of means as well as on his appreciation of the situation. The conditions of his action were such that the instrumental rationality guiding initially his behaviour made place to a form of axiological rationality. The actor changed in the course of the process.

2. GOOD INTENTIONS, NAÏVE EXPECTATIONS

In Fallada's novel, the deficit of information, although induced by the surrounding social system, is individual. It can also affect a collectivity. A group as a collective actor can make an unfortunate decision and provoke unanticipated consequences by an insufficient knowledge of the mechanisms of interdependence involved. Some decades ago, engineers of the NASA were looking for some food appropriate to long-lasting space missions. Their attention was drawn to a plant of the Andean regions called quinoa, which had been the staple diet of the Bolivian population for centuries and seemed to offer an excellent nutritional equilibrium. The direct consequence of the NASA discovery was the decision to make quinoa available to the cosmonauts. But the news of this discovery spread out. As a by-product of NASA activities, quinoa, little known outside Bolivia until then, became very much in demand in the United States and in Europe. With the help of development agencies and the encouragement of regional leaders, Bolivian farmers began to cultivate quinoa for exportation. The growth of production combined with an easier access to the international market contributed substantially to rural development and everyone was pleased with the success of the operation in its first stage. Very soon, however, an unintended consequence of this economic change became a matter of concern: the rise of the price of quinoa on the national market prevented more and more Bolivians to consume their healthy traditional food. They were obliged to turn massively to cheaper industrial food, increasing dramatically the risk of chronic malnutrition. At the moment, this unintended consequence of a well-meaning programme of economic development has become a large scale social problem calling for new public interventions (Romero and Shahriari, 2011).

The basic assumption behind the project was that an increase of revenues of the Bolivian farmers was bound to have repercussions on the whole economy of the country and lead to a general improvement of the living standard of the population. The consequences at the international level, however, were unanticipated. The demand for quinoa in the developed world increased more than was previously expected; soon, quinoa was advertised in American and European supermarkets as a

very healthy food, worth to be purchased, and became fashionable. A strongly export-oriented production generating higher costs but requiring new skills to make the product marketable substituted itself to the traditional small-scale production. At the root of this unanticipated chain of events was an insufficient identification of the various elements of the relevant system when the Bolivian market was considered in isolation. It appeared that the change in production patterns, although to the undeniable advantage of a certain segment of the population, did not benefited to the population as a whole. On the international market, the search for a healthy diet became predominant, which means that the improvement of the living standards of the Bolivian peasant population stopped to be of primary importance. The initial goal made room for diversified goals and new actors appeared on the scene.

3. CONFLICTING VALUES

Unintended consequences of a programme of action are very often the result of an inability to master the complexities of a social system, and that is why an incremental way of designing policies is often advocated (Dahrendorf, 1984: 255–258). There is no such thing as a rational policy *in abstracto*: the rationality of the policy maker is always a *bounded rationality:* a rationality relative to a certain frame of reference (March and Simon, 1958, 1979: 136). Every policy programme brings about a counter-programme taking gradually the action away from its initial goal. Voluntarist interventions, furthermore, can also reveal the existence of conflicting values. This can be illustrated by the current situation in Venezuela. Although Venezuela is one of the world most important oil producers in a time of soaring prices, this country is affected by shortages of basic products like milk, beef, chicken, vegetable oil and sugar. Surprisingly, such shortages affect both the poor and the well-off, even if the second ones are better equipped to cope with them. The queuing in front of the State sponsored stores subsidized by the government, before opening time is the most visible sign of a situation of chronic scarcity but the shortages are the result of a reduced production as well as of a disarray in the field of distribution. It seems that this situation arose when the authorities, in an effort to regulate transactions on the

market, imposed strict price controls. The rationale behind this state intervention was to make a series of essential goods affordable to the lowest classes of the population, given the extreme gap between rich and poor. Ironically, these essential goods were the hardest to find after the enforcement of tight regulations including for example required price cuts on fruit juice, toothpaste and disposable diapers. It is the usual dilemma of the promotion of equality opposed to the preservation of freedom of business and trade. Some economists argue that one has to deal here with a classic example "of government causing a problem rather than solving it" and indeed the exchange of arguments is reminiscent of situations of crises in other countries at other times. According to the authorities, some products are deliberately held off the market in order to push up the prices. And "profiteering companies" are designated as culprits. Critics of governmental interventions in a free market economy point out that when prices are set very low, companies and producers are no longer able to make a profit and react accordingly. And this is conducive to a chain of unintended and unanticipated consequences: farmers grow less food, manufacturers cut back production, retailers stock less inventory, products missing from retail shops show up on the black market at much higher prices (Neumann, 2012). This is an illustration of the fact that the more individual and collective actors involved in a process are numerous, the less the probability that a course of action unfolds according to the plan. In the case of Venezuela, aside from conflicting values there are conflicting interests and the evidence of a power struggle. For all its demagoguery, the populist Chavez regime has unquestionably taken a lot of measures in favour of the deprived categories of the population. For the very poor who could barely survive in other times, queuing to acquire essential goods is a bearable inconvenience compared with the possibility to afford them at all. For the authorities, the unintended consequences of a policy aiming at bridging the gap between rich and poor can be seen as just collateral damages.

4. THE INTERDEPENDENCE OF ACTORS

The very interdependence between collective actors can be as such a source of unanticipated consequences when there is a turn of events

modifying the rules of the interdependence. The transition between an administered economy and a market economy is a good example of such transformation. A few years ago, the Chinese authorities ordered the utility companies to freeze electricity prices. A seeming rational policy was thus initiated with a lack of foresight due to insufficient knowledge of the changing systemic context of the action. The governmental edict was promulgated to shield people from rising prices and help households to make ends meet. But in a context that was no longer that of an administered economy, the order had very soon unexpected consequences. Power companies reacted by cutting back on buying charcoal. Mines reduced production accordingly. Later on, winter weather required more charcoals than what was being produced and power plants were caught with very low stocks, resulting in blackouts and factory shutdowns (McDonald, 2008). Authorities reacted to inflationist tendencies with measures which would have been appropriate in the context of a planned economy but which were no longer so in an economy already dominated by market mechanisms.

More generally speaking, unforeseen events can add unintended consequences of decisions to the expected ones. Actions which can be considered adequate in order to help reaching a certain goal become a source of unintended consequences in a new conjuncture when the systemic environment is transformed. Mohamed Cherkaoui, in his book on Sahara, points out that France as a colonial power manoeuvred to enlarge the Algerian territory under its jurisdiction. This policy was to be explained by the fact that Algeria, unlike Tunisia and Morocco, could be placed under direct administration. Therefore, territories which were traditionally part of Morocco were incorporated into Algeria. Of course, in post-colonial time, that policy has completely lost its relevance. As a long term unintended consequence, when colonial time was over and Algeria became independent, the transferred territories became a source of conflict between Algeria and its neighbours. An ultimate consequence was that the military grew increasingly powerful in the political system of independent Algeria (Cherkaoui, 2007).

5. WISHFUL THINKING AND INADEQUATE ANALOGIES

The decision to create a common European currency—the Euro—to be substituted to the national ones appears afterwards as foolhardy. Nowadays, one wonders how the idea to form a currency union prior to a political union has been accepted ten years ago since it was clearly a way of putting the cart before the horse. It seems obvious that the precedent of the creation of the European Coal and Steel Community (ECSC) by the Paris Treaty in 1951 played a not insignificant role in the process of decision making. The political gamble of the fifties was taken with World War II still very much in the background. The plan—which was a political one—was to bring together the production of coal and steel of the member states to prevent the repetition of the events of the thirties when Germany, using these strategic important goods, developed stealthily a dangerous armament programme. But foremost in each participant's mind was the anticipation that economic cooperation would be the driving force to a political union. Nevertheless, the analogy between the creation of the ECSC and the creation of the Euro was misleading since the problems raised by a monetary union are completely different than those raised by a cooperation in the field of industrial production. Confronted by the economic and financial crisis, policy makers realize that the creation of the Euro in an environment shaped by national States jealously guarding their prerogatives was a "fair-weather" project and that a kind of federal authority becomes unavoidable if the Euro is to survive as a common currency. But institutional changes of this magnitude take time: in the 19th century, it took forty years to the Swiss Confederation—already a federal State in spite of the name—to achieve the Swiss monetary union.

However, pressed to react quicker than usually to avoid the break up of the Euro zone, policy makers tend to look for economic remedies short of a complete reorganisation of the union. And again, superficial generalisations based on inadequate comparisons lead them to stick to policies which have proven to be the source of unintended consequences in the past. Indeed, the current economic crisis offers a textbook example of such blindness. Quoting from John Maynard Keynes,

the Nobel Prizewinner Paul Krugman recently reminded us that "the boom, not the slump, is the right time for austerity" (Krugman, 2012a; 2012b). However, the teaching of Keynes is neglected nowadays by most of the decision-makers and their economic advisers. It is as if the lessons of the great depression of the thirties were completely forgotten. The failure of austerity policies to produce an economic recovery is due to the comparison of a national economy with a private household—a comparison which certainly appeals to common sense. This, according to Krugman, relies on "a bad metaphor". There are certainly good reasons for advising families to balance their budgets and to avoid accumulating debts. If debts start piling up, the sensible way out of the problems is for family members to "tighten their belts". But the comparison is not helpful because a national economy does not function like the domestic economy of a household. At the national level, the income mostly comes "from selling things to each other" which means that a person's spending is another person's income and vice versa. If everyone limits its spending at the same time in order to reduce debt, everyone's income will fall, especially when the public sector restricts its spending at the same time as the private sector. Generally speaking, the income of one category will fall because another category is spending less. If everybody's income plunge, the national debt problem gets worse, not better (Krugman, 2012a).

6. A SYSTEM IN TRANSITION

As an institution affected by a growing influx of students of various backgrounds, University is currently characterised by an uneasy competition between the traditional collegial model and a new bureaucratic model. It is a transitional system in which the delegations of authority are undefined, the responsibilities dispersed and the possibilities to resist a decision frequent. In such a complex organisational context, it is not uncommon to observe sequences of motivations and actions leading to unintended consequences. The illusions raised by the geographical mobility of students are a case in point. During the last twenty years, everywhere in Europe, academic authorities undertook to stimulate what came to be called "student mobility" at the international level.

The project which was to be called *Erasmus* was to encourage university students to go to a foreign university to take courses there during half a year. The idea was indisputably attractive. It was supposed to revive the tradition of the "travels to Italy" of the past, to improve the general culture of the students and to lower the psychological barriers between young Europeans of various nations. To make this ambitious programme possible and to avoid that students willing to take the risk of a hiatus in their studies were put at a disadvantage in their home university, a harmonization of procedures and programmes at the European level became necessary. This was an strong incentive to introduce a system of transferable study credits within the framework of ECTS (European Credit Transfer System) and to organize one semester courses and end-of-semester exams. At first glance, these reforms were entirely procedural and neutral from the point of view of the standard of high education and the specificity of the university institution. However, in a context which is far from being reminiscent of Erasmus's intellectual wanderings, the whole enterprise contributed to the deterioration of the high education system. When is was already too late to reverse course, it appeared that the new practical arrangements were in total contradiction with the spirit of the classical Humboldt University. When a course ends a few days before the corresponding examen, it could not be expected that the students make any effort to deepen their knowledge through a personal effort, for example by doing some additional reading. In fact, the quest for excellence makes place to the hunt for credits. As for the time spent in foreign universities, it is very often the occasion of bringing back home study credits obtained in a dubious way. Indeed, while being abroad, in a new environment, confronted with a new language, the foreign student is being given "the benefit of the doubt" when his performance is not convincing. In spite of its undeniable success among students, the Erasmus programme contributed to the transformation of universities in pure schools (Coenen-Huther, 2001: 131–132).

7. THE WORKING OF A POLITICAL INSTITUTION

The Swiss political system, based partly on "direct democracy", allows for groups of citizens to launch so-called "popular initiatives" that is

to collect signatures in order to force a vote on a project of constitutional amendment. A few years ago, a group led by the environmentalist Franz Weber managed to present successfully an "initiative" aiming at limiting the number of secondary residences in all municipalities. The proposal was to fix a ceiling at twenty percent of the total number of residences. Its initiators were aiming at preserving the landscape and stopping the increase of the number of residences occupied only a few weeks each year. It could reasonably be expected that a majority of the population would support such a proposal. And indeed a majority of the voters pronounced themselves in favour of it. But the majority was very slim—a little more than 50% of the voters - and concentrated in towns. In the Alpine regions where a great deal of the inhabitants live from tourism, the "initiative" was massively rejected. In the mountain areas were land ownership is very fragmented, there was a rush from small land owners to obtain from the local authorities permission to transform old barns or haylofts in residence units before the new legislation implementing the "popular initiative" would be brought into force (Chuard, 2012). On the other hand, property developers operating on a large scale had taken the precaution of asking for building permits long before the date of the vote. As an unintended consequence of this exercise in direct democracy, the race to get building permits has never been so intense in the touristy municipalities. Some figures are impressive: in Nendaz, canton of Valais, 141 requests have been introduced in 2012, during the three months following the vote, which is more than four times the average of the period 1994–2011, in Laax, canton of Graubunden, 252 requests of building permits were introduced during the same period, while the average of the previous years was 15 (Patrick Monnay, 2012). In short, an initiative aiming at solving a problem—to stop a sprawl of concrete to spoil the Alps—had the unanticipated consequence of making it acuter. The Weber initiative which was supposed to give concrete expression to values cherished by the majority of the population resulted in the humiliation of large segments of the population in alpine cantons and the creation of tensions between urban and rural areas.

8. CONCLUSION

Looking back to the examples summarized here, we realize that unintended and unanticipated consequences of actions are very much part of any social system. Using a Durkheimian formula, I would say that it has to do with the *sui generis* character of social interactions. Social life is permanently fluctuating. In the course of a process, the actors change, the system itself is transformed, the conditions of the action are modified by the interaction between actor and system. Quangel facing death after the failure of is action is no longer the man who started disseminating vengeful messages, hoping for a result. The Bolivian agriculture with access to the international market is no longer the system of agricultural production for local consumption. The students of the mass university do not behave the same way as those of the classical university. Actions produce unintended consequences because a social system, as a complex adaptative system, is a *human* system. The interaction between human beings include always an element of contingency which makes it partly unpredictable. The greater the number of actors involved, the greater is the contingency. And if actors are part of different groups, organisations or institutions, various logics of action are in competition. Jon Elster illustrates this with a very simple but illuminating example (1992: 9). He invites us to imagine that an individual is facing the eventuality of participating in a collective enterprise, be it a practical activity or, for example, an electoral campaign. He can reason in two different ways, each one being entirely plausible. Either he think that if most of the members of the collectivity are involved he has to participate as well, or he concludes that if the majority of the members participate there is no need for him to do so. Of course, we can understand both reasonings. As Hayek would say, we do so on the basis of an unconscious introspection (2010 [1952]: 123). We conceive of an observed or imagined behaviour by thinking about ourselves (*ibid.*: 104). Without additional information, however, we can imagine both reasonings but we are not able to foresee which one will be chosen. Indeed, we do not have at our disposal a theory encompassing both terms of the alternative. On the other hand, once the choice has been made, we can appreciate its logic even if we do not approve it. The first option could be

called the "civic one": the individual feels a moral obligation to join the others. The second option, favoured by rational choice theorists, would then be called the "asocial one"; the individual adopts a pure utilitarian attitude and does not bother to participate if he can avoid it. The decision-maker is in the same position as the sociological analyst; he cannot know for sure how the people affected by his plans will react. As Quangel did, he can project his own motivations on others but it is nothing more than a guess which can be erroneous. Different hierarchies of value, different cultural backgrounds can interfere with his intuition. Every decision to act, with the anticipations it implies, has the character of a psychological experiment. The sense of evidence regarding the consequences is a matter of routine more than a matter of reason. When the Americans invaded Iraq to topple Sadam Hussein, they made a series of strategical decisions that were supposed to be adequate for their purpose. Repeatedly, unintended consequences of their actions showed them that Iraq was not Germany 1945 and that their decision-makers had blundered. Unintended consequences of human actions are so much an unavoidable ingredient of social life that no sociology would be necessary if there was none. The paradoxes of the consequences are also the paradoxes of sociological analysis.

REFERENCES

Buckley, Walter (1967) *Sociology and Modern Systems Theory*. Englewood Cliffs, NJ: Prentice- Hall.

Cherkaoui, Mohamed (2006) *Le paradoxe des conséquences*. Geneva: Droz.

——— (2007) *Le Sahara, liens sociaux et enjeux géostratégiques*. Oxford: The Bardwell Press.

Chuard, Patrick (2012) "Résidences secondaires: c'est la course aux permis", *Tribune de Genève*, 16 August.

Coenen-Huther, Jacques (2001) "Les acteurs et les conséquences inattendues de leurs actions: le cas de l'Université", *Revue Européenne des Sciences Sociales*, XXXIX(121): 125–140.

——— (2012) *Les paradoxes de la sociologie*. Paris: L'Harmattan.

Crowe, Eyre (1928) "Memorandum on the present state of British relations with France and Germany", Foreign Office, January 1907. In G.P. Gooch and H. Temperley (eds), *British Documents on the Origins of the War*, Vol. 3, *The Testing of the Entente*. London, H.M. Stationery Service.

Dahrendorf, Ralf (1984) *Reisen nach innen und aussen*. Stuttgart: Deutsche Verlags-Anstalt.
Elster, Jon (1992) *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.
Ewing, Jack (2012) "ECB chief critical of euro zone's leadership", *International Herald Tribune*, June 1st.
Fallada, Hans (1965) *Jeder stirbt für sich allein*. Berlin: Aufbau-Verlag.
Hayek, Friedrich von (2010 [1952]) "Scientism and the Study of Society", in *Collected Works*, Vol 13, Part I. London: Routledge, pp. 75–166.
Krugman, Paul (2012a) "The austerity agenda", *International Herald Tribune*, June 2.
——— (2012b) *End the Depression Now!* New York: WW Norton.
McDonald, J. (2008) "Cooling inflation communist style", *International Herald Tribune*, February 12.
March, J.G. and Simon, H.A. (1958) *Organizations*. New York: John Wiley and Sons.
Monnay, Patrick (2012) "Une course aux permis de construire", *Tribune de Genève*, 16 August.
Neuman, William (2012) "Struggling for food in an oil-rich economy", *International Herald Tribune*, April 21–22.
Parsons, Talcott (1951) *The Social System*. London: Routledge and Kegan Paul.
Romero, Simon and Shahriari, Sara (2011) "Indigenous treasure out of Bolivian's reach", *International Herald Tribune*, March 21.

15

INSTRUMENTAL, AXIOLOGICAL RATIONALITY AND THE EXPLANATION OF NORMS

Cherkaoui's (and Boudon's) Critique of Rational Choice Theory and Its Ability to Explain Norm Emergence

Karl-Dieter Opp
University of Leipzig / University of Washington (Seattle)

The present paper takes up some claims made in Mohamed Cherkaoui's *Invisible Codes* (2005, numbers in parenthesis refer to this book). He provides a devastating critique of rational choice theory (RCT) and of James Coleman's theory of norm emergence. Furthermore, Cherkaoui claims that RCT in general is not able to explain norms, and especially not systems of norms. These claims are discussed in the present article. The conclusion is that both assertions are not tenable.

THE CRITIQUE OF RATIONAL CHOICE THEORY

A problem with the critique of RCT is that most critics do not specify what version they address. RCT is not a theory but a whole family of theories. It is extremely important to be clear about the version one is

concerned with because some of the objections against the theory may (and do) hold for some versions only and not for others. In order to assess Cherkaoui's critique, the first question thus is what version of RCT does he criticise. Only then can it be discussed to what extent his critique is in general acceptable or only for a certain version of RCT.

THE TWO VERSIONS OF RATIONAL CHOICE THEORY

I distinguish two versions of RCT: a narrow and a broad one. A *narrow version* is applied in neo-classical economics and often in game theory. Major assumptions are that only egoistic preferences matter, that objectively given (and not perceived) constraints influence behaviour and that a utility function is maximized that is based on the objectively given behavioural options and constraints.

A *broad version* grants that these assumptions may hold in certain situations. The narrow version is thus a special case of the broad version. However, the broad version differs from the narrow one in the following ways. (1) All kinds of preferences are admitted as explanatory variables. For example, actors may have altruistic preferences (i.e. they are intrinsically motivated to enhance the wellbeing of others). There may, but there need not be egoistic preferences. Furthermore, internalized norms, i.e. the intrinsic motivation to follow norms, may influence behaviour. (2) Perceptions of constraints, i.e. beliefs, are explanatory variables. These beliefs may be wrong, but as long as the actors think they are true they determine their behaviour. (3) Actors do what they think is best for them. This may not be the best course of action, from the point of view of an observer, but the actor does what is best from his or her point of view.[1]

There are several theories that are usually not regarded as rational choice theories, but they share its basic assumptions. The theory of cognitive dissonance (see, e.g., Festinger, 1957) assumes that certain constellations of cognitive elements cause psychic strain. In other words, certain sets of cognitive elements are costly. Such theories outline new situations and new kinds of costs and benefits and should therefore be classified as hypotheses of a broad version of RCT.

It is of utmost importance that in applying the broad version, the preferences and beliefs of the actors must be measured. This holds for the narrow version as well. If it is assumed that actors have egoistic preferences one wants to see the evidence for this claim. The techniques of empirical research provide a widely used set of measurement tools.

Sometimes the question is raised whether the broad version can still be regarded as rational choice theory. As a matter of fact, the central variables of both versions are identical. It is therefore justified to call the broad version rational choice theory.

REASONS, BELIEFS AND THE BROAD VERSION OF RATIONAL CHOICE THEORY

Cherkaoui's critique of RCT and the alternative he advances is heavily influenced by the work of Raymond Boudon (see, e.g., 81, 107–110). Therefore, the following discussion is also a discussion of Boudon's cognitivist model (Boudon, 1989) which he now calls the theory of ordinary rationality (e.g., 2012b: 17–18). Boudon has presented and discussed this theory in various publications. I will mainly refer to his most recent work.[2]

Although Cherkaoui never specifies which version of RCT he criticizes, the discussion clearly shows that his target is the narrow version, as the following quotations indicate.[3] RCT assumes instrumental rationality (6). It is concerned with expected results, i.e. advances a consequentialist-type reasoning (70). Norms are thus not part of RCT. Instead, action is based on beliefs that actors perceive as strong and is to be explained by value rationality or axiological rationality (70). This alternative is a reasoned action conception (55) as it is proposed by Raymond Boudon.

Cherkaoui illustrates his critique with the voting paradox (see Riker and Ordeshook, 1968) which he calls an enigma for RCT (80–81). Similarly, according to Boudon "the paradox of voting would suffice to show that RCT is unable to explain numerous familiar behaviours … there is no solution to the puzzle on voting within the RCT" (2012a: 7; see also 2003: 6–7). I take this example to show that Cherkaoui's (and Boudon's) claims do not hold for a broad version of RCT.

A prediction of RCT is, as Cherkaoui and many others point out, that nobody will participate in an election. The parties in an election offer packages of public goods. To be sure, citizens have preferences for certain packages of public goods. But the single citizen does not have any influence on which party comes to power. The reason is that a society is a large group where a single person cannot influence the outcome of the election. Therefore, whether a citizen votes or not has only a negligible impact on the results of the election. In addition, voting is costly. It takes time that could be used for other purposes. So there is no incentive to vote. The prediction thus is that nobody votes. The paradox is that this prediction is plainly wrong.

In order to show that this is not the prediction of a broad version let us first formalize the explanation by applying value expectancy theory (VET) which is also called utility theory. This is a widely used version of RCT in economics and social psychology (see, e.g., Riker and Ordeshook, 1973; Feather, 1982). Applying this theory to explain voting yields the following equation (1) and (2):

$$\text{ExpUtility}\,(\text{Vote}_i) = (p_{n-1} + p_i) \times U_i(\text{PublicGoodsProvision P}) - \text{Costs}_i \quad (1)$$

$$\text{ExpUtility}\,(\text{NotVote}_i) = p_{n-1} \times U_i(\text{PublicGoodsProvision P}) \quad (2)$$

When a person votes, his or her *expected utility* of voting depends, first of all, on his or her utility U_i of the public good. Its provision depends on the likelihood of participation of n others (p_{n-1}, i.e. the n members of the group minus the individual i) and of the likelihood that the individual's participation makes a difference (p_i). This additional probability p_i is close to zero in a large group so that p_i can be dropped from equation (1). In addition, the individual incurs a cost (Costs_i) when he or she votes. The overall expected utility from not voting is captured by equation (2). Because p_i is zero and voting is costly, the expected net utility of voting is lower than that of not voting. Because the behavioural alternative with the highest expected utility is chosen, and because equations (1) and (2) hold for all voters, the prediction is that nobody will vote.

What is wrong with this argument from the perspective of a broad RCT version? The major problem is that numerous incentives that *might* be relevant are left out of the equations, and that the equations

contain actual probabilities and not perceived probabilities (i.e. beliefs). Empirical research indicates that several additional factors influence electoral participation. These factors have relatively strong effects on voting so that the general prediction from the broad version is by no means that nobody will vote. Research has shown that the following factors instigate many citizens to participate in elections.

(1) *The ethics of voting.* Most of the citizens of democratic societies learn that it is a civic duty to go to the polls. If individuals have internalized this norm, there will be moral satisfaction or a good conscience in case of norm conformity (NormConf or NC) which has a positive utility (U_{NCi}). Normally, an individual is certain that this utility occurs when he or she votes ($p_{NCi} = 1$). If an individual does not vote, there is norm violation (NormViol or NV) which leads to a bad conscience which has a negative utility (U_{NVi}). The probability (p_{NVi}) is 1 because the actor knows for certain that norm violation results in a bad conscience.

The norm to vote may be so intense that it overrides all other incentives. One may call those voters, *moral voters*: they only vote for moral reasons. The previous equations can thus be extended:

$$EU\ (Vote_i) = (p_{n-1} + p_i) \times U_i(P) + p_{NCi} \times U_{NCi}(NormConf\ C) - Costs_i \quad (3)$$

$$EU\ (NotVote_i) = p_{n-1} \times U_i(P) + p_{NVi} \times U_{NVi}(NormViol\ V) \quad (4)$$

(2) *The cognitive illusion of being influential.* Equations (1) and (2) include objective probabilities. There are various empirical studies showing that individuals overestimate their personal influence (see, e.g., Finkel, Muller and Opp, 1989; Opp, 2001). In other words, p_i is often greater than zero. In this case, the product term for P is no longer zero and may become a dominant motivation to participate. Such false beliefs are also included in Boudon's cognitivist model (see, e.g., Boudon, 2003: 12–14).

This type of voter could be called the *instrumentally irrational voter*: voting is instrumental for providing public goods, but it is irrational in the sense that actual influence is misperceived.

(3) *Social rewards and punishments.* Voting may entail positive social rewards (or, equivalently, sanctions) from important others, whereas abstaining may result in punishment. There may be situational

social rewards at the polling station: one may expect to meet nice people. In equation (3) a product term for positive sanctions and the likelihood that they occur might be included. Expectations of others that one should or should not behave in a certain way could also be included: these expectations may vary in importance (i.e. utility) and expectancy (i.e. subjective probability) and can thus be included in the previous equations as another product term. One could imagine that there are *social voters* whose only reason for voting are social rewards.

There may be alternative groups who do not reward but punish electoral participation because this stabilizes the capitalist system. In this case, there would be negative rewards for participating and positive rewards for abstaining.

Boudon (e.g. 1998: 820) mentions these factors as auxiliary assumptions which are, according to Boudon, not satisfactory. It seems that auxiliary seems to mean ad hoc. However, this is not the case. The factors are *kinds of utilities and beliefs* and, thus, factors the broad version of RCT refers to. They are thus not added ad hoc in order to salvage RCT (1998: 821). The factors are also not ad hoc because, as was said before, their existence has to be ascertained empirically.

The previous discussion has the following *five implications* that are relevant for Cherkaoui's and Boudon's critique of RCT.

(1) *Utility theory as a broad version of rational choice theory.* Expected utility theory is a target of Cherkaoui's critique of RCT. However, beliefs as well as a wide array of *consequences* (including norms) can be included in the theory. The term consequences does not only refer to external consequences (such as perceived social sanctions in the case of voting), but also to internal states that are related to a given behavioural alternative. For example, a consequence of pollution may be that others get hurt (external consequence). But a consequence may also be that an action is in line with an internalized norm or with the expectations of others that one should behave in a certain way. Furthermore, an action could lead to a feeling of moral satisfaction. Such a broad definition of consequences is clearly implied by VET, as the previous equations show. These could be called internal consequences.

VET is related to the more general version presented before in the following way. *Utilities* in VET are any kinds of *preferences*. For

example, a person's altruistic motivation is the utility that one's action enhances the welfare of others. The constraints are the *perceived* behavioural consequences (i.e. beliefs). The assumption that the action with the highest expected utility is chosen means that actors choose what they think is best for them. There is thus utility maximization from the perspective of the actors.

(2) *Instrumental rationality and axiological rationality in the broad version of RCT.* As was shown before, there is no problem integrating norms into VET. However, Cherkaoui, referring to Boudon, disagrees (e.g. 81):

> As Raymond Boudon clearly shows (…), voting is among those actions guided by principles rather than the consequences they might produce. If I vote it is not because I am maximizing my utility (instrumental rationality), but because I think it's a duty and that elections are a good thing without which democracy, the best or least bad political regime, would be impossible (axiological rationality).

There are several problems with this argument. (a) Empirical research shows that voting is not only guided by principles but by external consequences as well (such as getting social rewards). (b) But let us assume there are only moral voters. VET implies that following the ethics of voting (i.e. acting according to principles) yields a higher subjective expected utility than not voting. Put differently, a moral voter maximizes utility from his or her perspective. One might feel uneasy with the term utility because this might denote some more profane motivation than acting on principles. However, what is meant by VET is that those individuals who are guided by principles do what they think is best for them. If this is denied: why does a person act on principles? If one thinks a certain action is right or is the thing one must do, this means that this is the best an actor can do in a given situation.

This is perhaps even more evident in situations with mixed motives. Assume a voter accepts the norm to vote and, in addition, has the option of spending a vacation with friends at the time of the election. Although

a principle is involved in the decision to vote, the person will weigh the costs and benefits and does what he or she thinks is best. Thus, there are situations where principles are involved and where clearly utility is maximized.

(c) The term *instrumental rationality* often denotes action that is chosen to attain goals (similarly Boudon, 2011: 35) or to achieve a congruence between means and ends (Boudon, 2009: 33). This implies that acting according to principles is instrumental as well. One may distinguish two kinds of instrumental behaviour (or rationality), one is behaviour to achieve non-normative goals (such as getting approval from friends), the other to achieve normative goals (such as conforming to the internalized norm of going to the polls). Both are kinds of instrumental behaviour. Why do we need a special term for axiological rationality referring to acting on principles? Theoretically, such a distinction is not necessary. It is even superfluous because the broad version of RCT can be applied to explain any kind of instrumental behaviour: the goals may refer to pursuing norm compliance or achieving other goals such as getting approval from others.

(3) *The broad version of RCT as a reasoned action conception.* VET or RCT usually do not use the term reason but they refer to the phenomena this term denotes. To illustrate, VET implies that actors who vote have good reasons to do so. For the moral voter, the main reason for going to the polls is his wish to follow an internalized norm (or to act on principles). The "instrumentally irrational" voter's reason to vote is to provide certain public goods. Thus, *reasons refer to the product terms of VET* or, in general, to the costs and benefits of a broad version of RCT.[4]

(4) *Are reasons causes?* Cherkaoui answers this question in the negative, without providing detailed arguments. Why couldn't we say that the internalized norms of an actor to go to the polls was a cause of his or her action? Reasons precede actions, and reasons are regularly associated with actions. These are the major characteristics of causes, and they clearly apply to reasons. Incidentally, reasons are clearly causes in Boudon's writings (see, e.g. Boudon, 2011: 37).

(5) *The explanation and effects of beliefs.* The subjective probabilities in the product terms of VET are the actor's beliefs. The critique of Cherkaoui that beliefs are not part of RCT is thus not acceptable.

Furthermore, nothing in the theory claims that these beliefs must be correct.

Cherkaoui argues, based on Boudon, that the most important deficiency of RCT is that it is not capable of explaining beliefs (107–108). This holds especially for beliefs that do not refer to consequences (107–108). X is good does not refer back to its eventual consequences. Cherkaoui then summarizes Boudon's major proposition to explain beliefs. The basic principle, which Boudon calls the *cognitive equilibrium principle* (Boudon, 2012b: 18), says

> that people believe that X is true, acceptable, good, legitimate, etc. as soon as they have the feeling that X rests upon a set of acceptable reasons.

The major proposition thus is that actions, beliefs and attitudes have a common cause: reasons held by the actors (108).

To what extent is the broad version of RCT compatible with this proposition? Let us look at the following examples: (1) Why does somebody stop smoking? Because the person *believes* that smoking leads to cancer and because the person *wishes* to remain healthy. (2) Why does a person think he or she *should* separate trash in the household? Because he or she *accepts a general norm* (a value) to protect the environment. (3) Why does a person *dislike* driving a car (i.e. has a negative attitude toward driving)? Because the person believes that driving contaminates the environment and this *should be avoided*. (4) Why does somebody *believe* that the communist party rules China? Because this is reported by various media and observers who have no reason to lie. The reports are the *reasons* for the belief.

Can these questions be answered by a broad version of RCT? Example 1 refers to a behaviour that is to be explained. The reasons are beliefs and preferences. This explanation is thus compatible with VET. Example 2 refers to explaining acceptance of a norm. The reason is acceptance of a more general norm or, as it is sometimes called, a value. We will discuss below in detail how RCT can grapple with question 2. Example 3 explains an attitude. The explanation is compatible with the attitude theory by Fishbein and Ajzen (see, e.g., 2009): an attitude

towards an object depends on the extent to which an actor believes that the object has other properties and the extent to which the actor values these other properties. It would be cognitively dissonant if an actor believes that an object has many positive features, but at the same time values this object negatively. Example 4 refers to beliefs—which, it is often argued, cannot be explained by RCT (Boudon, 2012a: 8–9). Not believing that the communist party rules in China would contradict credible reports. This would cause strong dissonance, which would be costly. Dissonance theory pertains to the broad version of RCT, as was argued before. This explanation is consistent with Boudon's explanation of scientific beliefs: they are believed because it is congruent with the regularities I observe in the real world (2012a: 8).

These explanations suggest that a broad version of RCT is compatible with a reasoned action approach. This holds if such a version also includes social psychological theories that share the assumption that human action (in a wide sense) is guided by costs and benefits.

(6) *Anomalies.* There are several empirical findings that are not consistent with RCT. Apparently, the target is a narrow version, although this is rarely made clear (see, for a summary, Kahneman, 2011). Based on the work of Boudon, Cherkaoui mentions two results in particular that I will address (107–108).

(a) It is often argued, also by Cherkaoui, that there are situations where a behaviour is enacted although the costs seem to be much higher than the benefits. There are two examples, not mentioned by Cherkaoui: vendettas and the rescuing of Jews during the Nazi regime. In both cases, the costs seem extremely high, at least at first sight. A detailed analysis of the situation shows, however, that the high-cost hypothesis is at least not as obvious as it seems to be (for details see Hamlin, 1991 on vendettas, and on rescuing Jews see Kroneberg *et al.*, 2010; Opp, 1997; Varese and Yaish, 2000). These cases suggest that after a detailed analysis of the data it may turn out that the actors' behaviour has lower costs or higher benefits than a superficial analysis indicates. A first intuitive judgment may thus lead astray.

(b) What is missing so far is a detailed analysis that answers the question as to what extent the anomalies are not only inconsistent with the narrow version but also with the broad version of RCT. Assume this

is the case. Because the cognitivist model is consistent with the broad version of RCT, the anomalies would refute both models. I suspect that, to be cautious, many anomalies do not refute the broad version. It would be of utmost importance to provide a detailed analysis to explore this question.

(7) *The explanatory power of RCT and of the cognitivist model.* The conclusion so far is that the cognitivist model does not contradict the broad version of RCT. This is in line with Manzo (2012: 36) who writes that the cognitivist model is "an extreme variant of broad rational choice theory." However, there is a difference. Let us look again at Boudon's basic principle (see above) claiming "that people believe that X is true, acceptable, good, legitimate, etc. as soon as they have the feeling that X rests upon a set of acceptable reasons" (Boudon, 2012b: 18). When we compare this principle with the detailed propositions of the broad RCT version (including VET and, e.g., dissonance or balance theory), RCT provides much more detailed guidelines in regard to what the conditions in the specific situation are that explain some phenomenon. In regard to explaining actions, for example, compare the principle and VET. Nothing is said in the principle about products of utilities and probabilities and about how to combine them. There is thus a difference in informative content.

Manzo's discussion of Boudon's theory (Manzo, 2012) suggests that various findings from other research traditions and theories can be utilized to add more specific assumptions to the theory and, I would add, to a broad version of RCT as well.

EXPLAINING NORMS AND SYSTEMS OF NORMS

Cherkaoui provides a devastating critique of James Coleman's theory of norm emergence (Coleman, 1990: chapters 10 and 11). Moreover, he claims that RCT in general may explain single norms or one norm at a time, but not a "constellation of integrated norms" (5, 82). It is one of the "enigmas" of RCT to explain the "determination of one norm by another" (80). Even the ability to explain single norms is limited: the "theory of instrumental rationality merely contributes to a particular perspective in sociological tradition's understanding of norm

emergence, namely the Hobbesian one" (6). This seems to mean that only the emergence of norms by human design can be explained because this is what Hobbes claims: citizens transfer rights to a government which means that norms are set.

Two questions must be distinguished in exploring the possibilities of explaining norms by RCT. One is to what extent Coleman's explanation is acceptable. Indeed, Coleman's explanation may be deficient, but RCT may nevertheless be capable of explaining norms. Thus, pointing to problems of Coleman's explanation does not imply that RCT in general is not suited to explaining norms. The second question is whether RCT in general can explain norms.

To be sure, Coleman mentions the issue of explaining systems of norms, but this was "beyond the scope of a treatment of the foundations of social systems" (Coleman, 1990: 265). This means that Coleman does not want to address it. It is incorrect when Cherkaoui writes that "Coleman acknowledges that rational choice theory can only explain the emergence of one norm at a time" (82). Even if Coleman would acknowledge this, it would be wrong, as the remainder of the article tries to show.

A major claim of Cherkaoui is that "utilitarian considerations with regard to the harmful or useful consequences of our actions are of no help in understanding the sphere of behaviour concerned with norms and values" (83). Such "consequences" are called "externalities." The claim thus is that externalities are irrelevant for norm emergence. We first will discuss this claim. Second, the question is addressed whether RCT can explain the influence of norms and values on new norms. Third, we explore whether RCT can explain systems of norms. Finally, we ask whether RCT can only explain the "Hobbesian" type of questions, as Cherkaoui argues, i.e. the emergence of norms by human design and not the spontaneous emergence of norms.

Before we turn to these questions it is necessary to specify what is understood by "norms", because this concept is often used in a vague way and with different meanings (for an overview see Opp, 2013a). According to Cherkaoui, a norm or rule "prescribes or recommends what should be done; it proscribes what must not be done.… It is external to individuals and is distinct from motives, or in other words from

the reasons that an actor has for acting or not acting. As such, it is really a macro phenomenon" (45). A norm is thus a shared external expectation that something should or should not be done. A norm does thus not refer to internalization. It is not clear whether a norm means *accepting* an expectation or only *expressing* an expectation. Both need not coincide. I may accept a norm not to jaywalk, but I may not communicate (i.e. express) this expectation to others.

Coleman does not give an "explicit definition" (1990: 243), but when he writes that norms "specify what actions are regarded by a set of persons as proper or correct, or improper or incorrect" (242), this actually seems to be a definition. It further seems that Coleman speaks of norms only if they are backed by sanctions. He does not say this explicitly, he only mentions an empirical hypothesis: "Norms are originally enforced by sanctions" (242). In his theory of norm emergence, "sanctioning" is the major variable. For Coleman, norms are also macro variables. However, they are aggregations of individual behaviour. For example, shared expectations means that a group of individuals express expectations.

In explaining norms, it seems useful to ask three questions: (1) When do people accept a norm? (2) When do people utter normative expectations? (3) When do people sanction others? It is not possible in a single paper to address each of these issues. I will concentrate on explaining *norm acceptance*. I assume that those who accept a norm usually have (and express) normative expectations.

THE RELEVANCE OF EXTERNALITIES

Let us return to Cherkaoui's claim that externalities "are of no help in understanding the sphere of behaviour concerned with norms and values." In contrast, Coleman's major variable in explaining norms is externalities.[5] These are understood as behaviours of actors that are beneficial or costly to others. For example, if factories contaminate the environment they impose a negative externality. If a supermarket opens near a residential area, the prices of property will rise. This creates a positive externality for the residents.

Are externalities really irrelevant for norm emergence? There are many examples illustrating that externalities are one factor for the

emergence of norms. For example, pollution is one condition for legislation to protect the environment. Laws against smoking are also due to the negative effects of smoking on health. Externalities are also relevant for spontaneous norm emergence, as will be shown below.

It is important to note that the externality explanation is not functionalist. It is not argued that a consequence of a behaviour leads to its regulation, but that a *perceived* consequence enters the decision calculus of actors.

Actors may further *create* positive externalities by setting norms. Welfare legislation is an example: the goal is to promote the welfare of certain societal groups. Another example is the development of special languages in groups. One goal of the members is not to be understood by others which is regarded as a positive externality by the group members.

There are also cases where externalities are irrelevant or of secondary importance. Governments may issue laws to curb environmental pollution *only in order to please their constituency* without being intrinsically interested in removing the externalities.

These cases can easily be integrated into a rational choice framework. In terms of value expectancy theory, externalities may be perceived behavioural consequences. For example, assume the decision makers of a firm are aware of imposing costs on others by pollution. The reactions of others will be behavioural consequences of certain actions such as reducing pollution, denying that there is pollution, doing nothing, etc.

Thus, the general claim that externalities are irrelevant for explaining norms is clearly not tenable. Obviously, other factors are relevant as well, as the extensive RCT literature shows.

VALUES AS DETERMINANTS OF NEW NORMS

Another factor for explaining norms are existing norms and values. Cherkaoui emphasizes, with Max Weber, religious beliefs as conditions for the emergence of norms. More generally, values (i.e. general norms) may have the effect that a behaviour is regarded as "unjust" because it is not consistent with an existing norm. Examples for such cases abound.

One of the justifications for tax laws that aim at income redistribution is to change an "unjust" income distribution. The norm here is that only certain distributions are normatively acceptable. Conscription in the US at the end of the nineteenth century was regarded as "just": there was increased immigration and many native-born American citizens believed that the new immigrants "should not be left free to pursue their own interests. The majority demanded conformity with Anglo-American values and viewed citizen service as an obligation" (Perri, 2013: 436). The general norm was that immigrants should contribute to the provision of public goods in the United States. A law that prescribed conscription of immigrants was a means to realize this value. The constitutional court in Germany examines to what extent laws conform to the constitution. The examples illustrate that norms may be in a *hierarchical* relation.

The examples further illustrate two cases: one is that a given norm is not consistent with another *norm*; the second is that a given norm is not in line with a *behaviour.*

Can RCT theory explain what happens when norms are related to each other? Such hierarchical relations have only consequences for actors if they are *cognitive*, i.e. perceived (the following is based on Opp, 1983: 120–129). If two norms are cognitively inconsistent their simultaneous existence is costly to actors (or, put differently, causes psychic strain). For example, if the constitutional court rules that a given law is unconstitutional this is a cost to the lawmakers who are expected to change the specific law. Cognitive theories can be applied to predict what the consequences of such cognitive inconsistencies are. There may be logical relations between norms, but often the norms are not formulated clearly enough so that it is difficult or impossible to decide what their logical relation is. In order to explain what happens if several norms exist it is important to what extent they are perceived as consistent.

EXPLAINING SYSTEMS OF NORMS

Norms are sometimes classified according to the kinds of activities that are regulated. For example, labour legislation is a set (or system) of norms that is related to labour. Environmental law regulates

activities that influence the environment. Inheritance law specifies what should happen when people die. Property rights refer to norms about who is allowed to do what with various kinds of objects. Canon law refers to what an organization or a set of organizations—the church—is or are allowed to do. These examples illustrate that there are sets of norms (or laws) that have a common name and that are perceived as belonging together.

How can we explain such systems of norms? It seems that there are different mechanisms. One is that there exists a general *value* that contributes to the origin of a whole system of specific norms. Assume there is a widely accepted *value* (and, thus, a goal) to protect the environment. Assume further that there is a set of *beliefs* about which actions lead to the realization of this goal and which actions compromise it. These beliefs may be wrong, but they will nevertheless govern behaviour. This value and the beliefs are a first step in the origin of a system of laws and norms to protect the environment. First of all, governments or political actors may address environmental concerns in the population and initiate laws in order to remain in power (as public choice theory would argue). Further, interest groups may put pressure on governmental agencies to issue laws referring to protecting the environment.

There may also be unplanned norm emergence. Citizens who observe that others endanger the environment will sanction them. Depending on the intensity of the interest and the number of citizens who share this interest, sanctioning will increase to an extent that the sanctioned behaviour decreases in frequency and that a norm not to enact these behaviours will emerge.

There may be more than a single value that is relevant for norm emergence. In regard to labour legislation, there is a value (and an interest) in "just" wages, in "good" working conditions and in "good" pensions after retirement. These different interests will result in the processes described in the previous paragraph. The outcome will be systems of norms.

These examples suggest that RCT can be applied to explain systems of norms: values and interests, together with beliefs about what actions or measures realise these values and interests are conditions for various processes that lead to laws or norms.

The claim that RCT can explain systems of norms is confirmed by an extensive literature on "new institutionalism" in the social sciences (see the overviews in Rhodes, Binder and Rockman, 2006). One of these institutionalisms is rational choice institutionalism (for an overview see Shepsle, 2006). The contributions to this field provide various explanations of systems of norms. Furthermore, the institutional economics research programme (see, e.g., Furubotn and Richter, 2005) addresses the origin of systems of norms as well. Finally, economic history is concerned with the emergence of institutions such as property rights during world history (see, e.g., the work of Nobel prize winner Douglass North such as North and Thomas, 1973 and more recently North, Wallis and Weingart, 2009).

REASONS, AXIOLOGICAL RATIONALITY AND NORM EMERGENCE. NOTES ON BOUDON'S AND CHERKAOUI'S THEORY

Even if RCT is able to explain norms, there may be better alternatives. One could be Boudon's cognitivist model that is also in line with Cherkaoui's theoretical arguments. In this section the question is addressed whether Boudon's "basic principle" is compatible with the explanation of norms by RCT.

Boudon's theory says, as was mentioned before, that norms originate if there are good reasons for accepting a norm. What are good reasons? In explaining a specific norm Boudon's procedure is to reconstruct a deductive argument that implies the respective norm (or, more precisely, normative statement) as a logical conclusion. The premises of the argument consist of empirical as well as normative statements because a normative statement can never be deduced from empirical statements alone. "Then, if no alternative set of acceptable and mutually compatible statements leading to a different or opposite normative or axiological conclusion is available to an actor, he will consider that it is axiologically rational for him to assume the given normative or axiological conclusion is right" (2012a: 9). Note that norm acceptance is explained. Note further, that the previous citation takes the perspective of the actor. Third, apparently the actor maximizes utility because he or she will accept the "best" argument.

Boudon (2012a: 10–15—the following numbers in parentheses refer to this article; see also Boudon, 2012b: 23–26 and other publications) illustrates his theory with an example. This example will be presented in detail because it illustrates very well how Boudon's theory is related to RCT.

The example is taken from Adam Smith's *An Inquiry into the Nature and Causes of the Wealth of Nations* (1776). There was a "strong collective feeling among 18th century Englishmen … that miners should be well-paid, while soldiers should be satisfied with mediocre salaries" (10). Why should miners get a higher salary than soldiers? Boudon first reports several *empirical statements* that were true in the historical situation. Some statements describe properties that miners and soldiers had in common. For example, both had invested similar time and effort to acquire the skills and knowledge needed to do their work. The risks were similar too: miners as well as soldiers could be killed. A major difference is that a miner's death is an accident, whereas a soldier's death is a sacrifice for their country. This difference leads to strong symbolic rewards for the soldier "in terms of moral prestige, symbolic distinctions, glory, notably when he has won a battle, or funeral honours in the case of death on the battlefield" (11). Miners do not receive these rewards. Why are these differences relevant for the different salaries? "[T]he salaries of the miners should be higher; otherwise, an unjustifiable disequilibrium between the contributions and the rewards of the two categories would appear" (11). In other words, "equal rewards should correspond to equal contributions" (10). These are *normative statements*.

Thus, the contributions of miners and soldiers are equal, but soldiers get higher rewards (empirical statements). This should be compensated by higher salaries of the miners (normative statement). This is the explanation of the norm that miners should get higher salaries than soldiers.

At first glance, this explanation has nothing to do with RCT. However, a closer analysis of Boudon's argument shows that the example amounts to explaining the origin of a norm ("miners should get a higher salary than soldiers") by means of a more general norm, i.e. a value ("contributions should be equal to rewards")—a case that has been discussed before. To apply this value to specific situations, e.g., to

explain the different salaries of groups, it must be empirically determined what the contributions of the groups to society are and what their rewards are. If the empirical statements mentioned before are true, it follows *logically* that those who accept these statements will accept the norm that miners should get a higher salary than soldiers. Thus, the premises are the "reasons" for the norm to be explained which is the conclusion of the argument.

It is well known that often actors to not observe the rules of logic or mathematics, as especially the work of D. Kahneman and A. Tversky shows. So the question is why do individuals accept the logical conclusion? In our example, why does the common man accept the norm that miners should get a higher salary than soldiers, if he accepts the premises of the deductive argument? Why does a citizen not say: "Alright, Boudon and Adam Smith are right, there are more symbolic rewards to soldiers than to miners, and the value of their contributions should match the rewards, *but* I nonetheless feel that soldiers and not miners should get a higher payment." It seems that Boudon implicitly assumes that the conclusion is only accepted because otherwise there would be *cognitive dissonance*, which is costly. This may be so obvious in the present example that it is not mentioned, but it is implicitly assumed.

This implicit assumption is also suggested by the second component of the definition of "axiological rationality" mentioned before: there is no alternative set of statements "which might be preferable" (2011: 34). This implies that the actor chooses the argument that is best from his or her perspective. Only in this case is there "axiological rationality."

Boudon argues that the reasons for accepting the normative conclusion "are not necessarily instrumental" (13). He confronts this "non-instrumental" approach with an "instrumental" approach. We can agree with Boudon that there is nothing "instrumental" in a purely logical argument. However, when we add the neglected "psycho-logical" link between the premises and the conclusion, the conclusion is actually accepted because otherwise dissonance or psychic strain would occur. Accepting the conclusion is thus "instrumental" in the sense that the actor wants to achieve certain goals such as bringing his or her cognitive elements to a consonant state. This is consistent with a broad rational choice approach.

"Instrumental" could also mean that the beliefs do not refer to consequences (see, e.g., Boudon, 2003: 9; see also the discussion by Esser, 2004, and the discussion above in this paper about the meaning of "instrumental"). It is difficult to see why a theory should restrict the reasons actors use in a deductive argument. For example, a reason for paying higher salaries to miners than to soldiers could be the societal consequences of such a difference: higher salaries could attract more people to get jobs as miners which might be urgently needed by the mining companies. It seems that Boudon relaxes this restriction in other publications (e.g. Boudon, 2009: 36).

Boudon emphasizes several points that seem important in the present context. One is that actors need not be aware of the reasons (i.e. of the deductive arguments) that lead them to accept a certain norm. The task for empirical research then is to reconstruct the arguments of the actors, which is in line with what Boudon did. Incidentally, the procedure of finding "cognitive maps" could be applied (see Axelrod, 1976; Gallhofer and Saris, 1996). A second point that Boudon emphasizes is that the reasons might be different at different times. This suggests a third point that Boudon mentions: it is important that the statements the deductive argument consists of are empirically valid. Without providing empirical evidence the explanations would be ad hoc. This holds for applying the broad version of RCT as well, as was emphasized before.

In RCT explanations of norms, externalities play a major role. Are they consistent with Boudon's "principle"? One might say that strong externalities are "good reasons" for accepting norms to curb or eliminate externalities. In sum, then, I do not see that Boudon's explanation of norms and explanations of a broad version of RCT are inconsistent.

SPONTANEOUS NORM EMERGENCE AND THE ORIGIN OF NORMS BY HUMAN DESIGN

The previous sections do not address the question of when norms emerge spontaneously and when by human design. Cherkaoui and Boudon do not provide detailed hypotheses specifying when norms emerge in which way. Coleman's theory does not deal with this question either. In this section, I will illustrate one mechanism of spontaneous

norm emergence in order to show how RCT addresses the explanation of spontaneous norm emergence (for the following see Opp, 2002). Since the beginning of the 1950s research has shown that active smoking and being exposed to smokers increases the risk of lung cancer. This and the motivation to avoid these risks, are empirical facts. There is also a general norm that persons should not harm others. These facts (which could be formalized as a deductive argument) led to accepting a norm that one should not smoke in the presence of others, i.e. a norm not to expose others to negative externalities. Furthermore, there are metanorms that justify or demand sanctioning those who harm others. The more widespread sanctioning becomes, the more costly smoking becomes and the more an antismoking norm spreads. This spontaneous norm emergence was further supported by anti-smoking laws (based on pressure on politicians and on the perceived opportunity for politicians to gain support). This explanation sketch illustrates how spontaneous norm emergence and norm emergence by human design can be explained by a broad version of RCT. Furthermore, Boudon's "basic principle" can be applied here as well.

CONCLUDING NOTE

The major claim of this paper is that the cognitivist model or, to use Boudon's most recent term, the "theory of ordinary rationality", also advanced by Cherkaoui, is consistent with a broad version of the theory of rational action. This is actually in line with Boudon when we look at the title of one of his first papers about his theory: "The 'cognitivist model.' a generalized 'rational-choice model'" (1989). The target of Cherkaoui's and Boudon's critique is a narrow version of the theory that is increasingly regarded as outdated. The compatibility of the different theories is mainly obscured because of different vocabularies that actually have the same meanings. For example, value expectancy theory consists of products of utilities and subjective probabilities, which are "reasons" in terms of the cognitivist model. It is further argued in this paper, that a broad rational choice approach is well suited to explaining norms, systems of norms and to explaining when norms emerge spontaneously or by human design.

If it is granted that the two sets of theoretical propositions are not contradictory, it was argued that the explanatory content of the hypotheses of the broad version of rational choice theory specifies the relevant explanatory variables in much more detail than the cognitivist model. Consequently, the explanatory content of the former is greater than that of the cognitivist model. A more detailed analysis of the inter-theoretic relationships between the different theories would be useful.

NOTES

1. For a detailed exposition and discussion of these versions see Opp (1999) and Kroneberg and Kalter (2012) (with further references). The role of norms in RCT is analysed in Opp (2013b).
2. Early expositions of the cognitivist model are Boudon (1989; 1996; 1998a; 1998b). See also Boudon (2003; 2009; 2010; 2011; 2012a and 2012b). The most recent version is Boudon (2012a) which is an unpublished manuscript. I am grateful to Boudon that he sent me this manuscript.
3. This concurs with Boudon's critique of RCT. He asserts that "rational choice theory introduces the idea that the reasons moving social actors are basically egoistic" (2012a: 17). He further asserts that RCT introduces the fiction of a solipsistic *homo sociologicus,* whereas Boudon's cognitivist model recognises *homo sociologicus* as a social being. This critique is clearly directed towards the narrow version.
4. Note that Fishbein and Ajzen (e.g., 2009) denote one version of their theory the theory of reasoned action. This includes VET.
5. Incidentally, the externality explanation is originally not by Coleman but by Demsetz (1967).

REFERENCES

Axelrod, Robert (ed.) (1976) *Structure of Decision: The Cognitive Map of Political Elites*. Princeton: Princeton University Press.

——— (1986) "An evolutionary approach to norms", *American Political Science Review*, 80(4): 1095–1111.

Boudon, Raymond (1989) "Subjective rationality and the explanation of social behavior", *Rationality and Society*, 1(2): 173–196.

——— (1996) "The 'cognitivist model.' a generalized 'rational-choice-model'", *Rationality and Society*, 8(2): 123–150.

——— (1998a) "Limitations of rational choice theory", *American Journal of Sociology*, 104(3): 817–828.

——— (1998b) "Social mechanisms without black boxes", in Peter Hedström and Richard Swedberg (eds), *Social Mechanisms. An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press, pp. 172–203.
——— (2003) "Beyond rational choice theory", *Annual Review of Sociology*, 29:1–21.
——— (2009) "On axiological rationality", in Paul Hill, Frank Kalter, Johannes Kopp, Clemens Kroneberg, and Rainer Schnell (eds), *Hartmut Essers Erklärende Soziologie. Kontroversen und Perspektiven*. Frankfurt and New York: Campus, pp. 31–52.
——— (2010) "The cognitive approach to morality", in Steven Hitlin and Stephen Vaisey (eds), *Handbook of the Sociology of Morality*. New York: Springer, pp. 15–33.
——— (2011) "Ordinary rationality: The core of analytical sociology", in Pierre Demeulenaere (ed.) *Analytical Sociology and Social Mechanisms*. Cambridge: Cambridge University Press, pp. 33–49.
——— (2012a) "The micro–macro link. Notes on a planned course on "truth, beliefs, interpretation", unpublished manuscript.
——— (2012b) "'Analytical sociology' and the explanation of beliefs", *European Journal of Social Sciences*, 50(2): 7–34.
Cherkaoui, Mohamed (2005) *Invisible Codes: Essays on Generative Mechanisms*. Oxford: Oxford University Press.
Coleman, James S. (1990) *Foundations of Social Theory*. Cambridge, Mass., and London: Belknap Press of Harvard University Press.
Demsetz, Harold (1967) "Toward a theory of property rights", *American Economic Review* 57(2): 347–359.
Esser, Hartmut (2004) "Wertrationalität", in Andreas Diekmann and Thomas Voss (eds), *Rational-Choice-Theorie in den Sozialwissenschaften. Anwendungen und Probleme*. Munich: R. Oldenbourg, pp. 97–112.
Feather, Norman T. (1982) *Expectations and Actions: Expectancy-Value Models in Psychology*. Hillsdale, NJ: Lawrence Erlbaum.
Festinger, Leon (1957) *A Theory of Cognitive Dissonance*. Stanford: Stanford University Press.
Finkel, Steven E., Edward N. Muller, and Karl-Dieter Opp (1989) "Personal influence, collective rationality, and mass political action", *American Political Science Review*, 83(4): 885–903.
Fishbein, Martin, and Icek Ajzen (2009) *Predicting and Changing Behavior. The Reasoned Action Approach*. New York and Hove: Psychology Press.
Furubotn, Eirik G., and Rudolf Richter (2005) *Institutions and Economic Theory: The Contribution of the New Institutional Economics*, 2nd edn. Ann Arbor: University of Michigan Press.
Gallhofer, Irmtraud N., and Willem Saris (1996) *Foreign Policy Decision-Making. A Qualitative and Quantitative Analysis of Political Argumentation*. Westport: Praeger.
Hamlin, Alan P. (1991) "Rational revenge", *Ethics* 101(2): 374–381.
Kahneman, Daniel (2011) *Thinking, Fast and Slow*. London: Allen Lanew.
Kroneberg, Clemens, Isolde Heintze, and Guido Mehlkop (2010) "The interplay of moral norms and instrumental incentives in crime causation", *Criminology*, 48(1): 259–294.

Kroneberg, Clemens, and Frank Kalter (2012) "Rational choice theory and empirical research: Methodological and theoretical contributions in Europe", *Annual Review of Sociology* 38: 73–92.

Manzo, Gianluca (2012) "Reason-based explanations and analytical sociology", *European Journal of Social Sciences*, 50(2): 35–66.

North, Douglass C. and Robert Paul Thomas (1973) *The Rise of the Western World. A New Economic History*. Cambridge: Cambridge University Press.

North, Douglass C., Wallis, John Joseph and Weingast Barry R. (2009) *Violence and Social Orders: A Conceptual Framework for Interpreting Recorded Human History*. Cambridge: Cambridge University Press.

Opp, Karl-Dieter (1983) *Die Entstehung sozialer Normen. Ein Integrationsversuch soziologischer, sozialpsychologischer und ökonomischer Erklärungen*. Tübingen: Mohr Siebeck.

——— (1997) "Can identity theory better explain the rescue of Jews in Nazi Europe than rational actor theory?", *Research in Social Movements, Conflict and Change*, 20: 223–253.

——— (1999) "Contending conceptions of the theory of rational action", *Journal of Theoretical Politics*, 11(2): 171–202.

——— (2001) "Why do people vote? The cognitive illusion proposition and its test", *Kyklos*, 54(2/3): 355–378.

——— (2002) "When do norms emerge by human design and when by the unintended consequences of human action? The example of the no-smoking norm", *Rationality & Society*, 14(2): 131–158.

——— (2013a) "Norms", in James D. Wright (ed.), *The International Encyclopedia of Social and Behavioral Sciences*, 2nd edition. Oxford: Elsevier.

——— (2013b) "Norms and rationality. Is moral behavior a form of rational action?" *Theory & Decision*, 74(3): 383–409.

Perri, Timothy J. (2013) "The evolution of military conscription in the United States", *The Independent Review*, 17(3): 429–439.

Rhodes, R.A.W., Sarah A. Binder, and Bert A. Rockman (eds) (2006) *The Oxford Handbook of Political Institutions*. Oxford: Oxford University Press.

Riker, William H., and Peter C. Ordeshook (1968) "A theory of the calculus of voting", *American Political Science Review*, 65: 25–42.

——— (1973). *An Introduction to Positive Political Theory*. Englewood Cliffs, N.J.: Prentice Hall.

Shepsle, Kenneth A. (2006) "Rational choice institutionalism", in R.A.W. Rhodes, Sarah A. Binder, and Bert A. Rockman (eds), *The Oxford Handbook of Political Institutions*. Oxford: Oxford University Press, pp. 23–38.

16

SOCIAL INTERDEPENDENCE, RIGHTS AND NORMATIVE DYNAMICS[1]

Emmanuel Picavet
Université Paris-1 Panthéon-Sorbonne

1. INTRODUCTION

In this article, I address the interplay between rights and freedoms on the one hand and the reasons of rights-exercising in a social context on the other hand. Presumably, this problematic should give a role to the positive or negative welfare effects of rights-exercising on other people in the context of a social interaction which is regulated by evolving social norms. After all, rights are social relationships and, building on antecedent insights from Max Weber's sociology, Mohamed Cherkaoui has been able to clarify the precise meaning of the notion that interdependence is to be hold predominant in social facts (Cherkaoui, 2006, esp. ch. 4). A core component of this achievement of sociology is the ability of human agents reciprocally adapt to the actions of one another. In the case of rights, this strongly suggests that the values involved in the actions which qualify as the exercising of rights should be considered dependent on the choices and underlying values of other people. A second theme which is of much interest to sociologists and philosophers alike is the interdependence of individuals through socially enforced norms and collective institutions. As emphasised by M. Cherkaoui

(2006: ch. 9), commenting Coleman's theory in a series of detailed and clear analyses, individual interests play a major role in the dynamics of institutions and norms. Starting from this, we could expect the normative and institutional relationships which express and implement "rights" to be found dependent on the interplay of (evolving) linkages between individual interests in social life.

But individual rights, especially those which are hold to be expressive of "freedom" in a moral sense, give *prima facie* reasons for giving little or no weight (from the viewpoint of a deliberating agent) to the welfare or values of the potentially affected parties. Indeed, rights-exercising seems to be associated with disregard for the affected parties when it comes to assessing the moral acceptability of acts. Given that people are the holders of moral rights, or the holders of rights in a legal sense provided these are part of a morally admissible legal system, further examination is not required to determine whether it is acceptable of not that they act in a given manner, provided that their so acting qualifies as a way to exercise their rights. Isn't this part of the meaning of rights, after all? This property illustrates the strong interdependence between legal rules (which give objective validity to rights) and a predominantly deontological conception of morality (which gives priority to normative injunctions over the comparison of possible results associated with alternative courses of action).

But thanks to Sen's pioneering contribution (1970a; 1970b), reconciling a concern for rights and respect for unanimity among the affected parties (if there is a unanimous agreement) is known to be logically problematic in social choice occasions, even though both principles are grounded in individualistic ethics. Sen's logical problem prevents respect for minimally specified rights-exercising prerogatives from being associated in a systematic way with consensual agreement among the affected parties. This is a possible motivation for the libertarian defence of radical views, according to which individual rights and freedoms should be viewed as side constraints which help delineate the domain of possible collective choices. Rights and freedoms, it is hoped, would then be insulated from welfare considerations.

I will try to establish that this view of rights cannot always yield a plausible account of rights-exercising. In many cases, individual rights

are such that the way to exercise them in a social context appears basically defective (rather than impolite in the context) if it does not rely on a significant degree of attention to the affected parties (sec. 2). This gives credibility to a working hypothesis: patterns of interaction, especially through the level of interdependence they express, give clues for the understanding of individual reasons in rights-exercising (sec. 3).

2. RIGHTS-EXERCISING AND THE INTERESTS OF THE AFFECTED PARTIES

a. Relevance of the central controversy about the description of rights

The moral priority of rights is classically expressed by the familiar thesis according to which abstaining from the infringement of rights and doing one's duties is more fully compulsory, or is more strictly called for, than promoting welfare or acting out of respect for various interests or concerns of the affected parties. In many avenues of social life, individual welfare and the values of individuals are quite simply screened off by rights and duties in the informational basis of decision-making. This plays a role, presumably, in the special connection between individual rights and legal rules. In a rough model or rights-holding and rights-exercising, a necessary and sufficient requirement for respecting the rights of everybody is to abide by the rules which enshrine those rights. What is implied is neither charity nor a sympathy for the affected parties which would go beyond legal obligations.

As shown by Sen's famous "liberal paradox" (Sen, 1970a; 1970b), even minimally specified individual rights, when they are strictly protected at the social level, conflict with the equally familiar intuition that collective choices are justified if they are consensual at some level in the stratified layers of social life (respect for unanimity or "Pareto principle" in pairwise comparison of alternative states of affairs). There is a conflict as long as we insist that collective decision-making should be minimally "rational", by which is meant the lack of evaluative cycles.

Individuals should have rights and social life should respect these rights: on the face of it, this statement shows no discrepancy with the notion that social decisions generally speaking are admissible if all the

concerned members of the considered society give their individual consent some way or other (if not personally, through their representatives and if not for substantial reasons—in case they disagree—at least indirectly, out of respect for the general pattern of well-justified procedures). Normative individualism is at the root of both claims. Respect for the autonomy of persons is expected to imply respect for fundamental human rights and a "private" domain of life; in addition, respect for individual will or freedom of choice provides a sufficient reason to let one have it one's way, provided there is no disagreement among those who might disagree. But the mutual compatibility of such benchmark expectations is highly problematic, given that individual preferences are left unconstrained (as shown by A. Sen).

In Sen's original contribution, "liberalism" or "minimal liberalism" was not put forward as an adequate description of freedom in any sense of the word. Rather, it was portrayed as some sort of minimal condition any conception of freedom should include or imply. Non-compliance with such minimal "liberalism" was therefore perceived as synonymous with radical oppression. Indeed, "liberalism" was identified with the ability of individuals to be decisive on pairwise choices between alternative social states of affairs (their preferred option is the one which will prevail). Given the standard view of "social states" as states of affairs which are characterised by the description of everything of interest to at least one member of society, liberalism in this sense is not a demanding condition: highly personal, socially irrelevant characteristics of the world can be found in the more "personal" regions of life. By the way, Sen's impossibility result applies even to "minimal liberalism" (with personal two-way decisiveness required for two persons at least).

Whatever the details of our conception of freedom, it should be incapable of being applied to a situation in which such a weak requirement of "liberalism" is violated. Accordingly, Sen-type liberalism and the "minimal liberalism" variant were described as plausible necessary conditions for *any* social choice procedure, provided that we accept the analytical format set forth by Arrow (1951).

This interpretation of the paradox, however, can been challenged. The emergence of rival systematic views of rights, notably in (broadly) game-theoretic approaches, has fuelled the controversy about the "true"

nature and destination of rights. Some authors have observed that the very meaning of rights is, as revealed by the specialised debates, affected by uncertainties (Brunel and Salles, 1998). Given the ontological controversies about their precise expression or description (as explored by Pétron-Brunel, 1998, Igersheim, 2004), rights can be called essentially contested concepts.

One of the effective lines of criticism has come from philosophers who, following Nozick (1974) and occasionally building on the technical analyses of Gärdenfors (1981), have claimed that rights are best understood as framework arrangements which should be considered logically prior to, and independent from, the social-choice procedures for amalgamating preferences or values. Rights may not be portrayed as properties of social procedures, nor of social life at large. Given such assumptions, individual rights-exercising can hardly be based on the details of social interaction; neither can it flow from the attention to the values and interests of other people. It must be understood as a decentralised device for setting the framework of future social interaction. Rights-exercising, whatever the content of rights, then becomes an individual-centred affair, and the natural assumption to retain is that rights-exercising reflects individual preferences (since "preferences" in social-choice parlance are a synthesis of every reason of choice or judgment for individuals).

Thus, rights could easily be described as belonging to individuals and some libertarians are prepared to see them as constitutive of one's personality or individuality. The true nature of rights-exercising could be equated with enabling individuals to determine some features of the world. It can safely be said that this view of rights has been prominent in the recent development of action-centred or game-theoretic analyses of rights systems (in line with such contributions as Sugden (1985), Gaertner *et al.* (1992). Can such a vision of rights reinforce the natural-right beliefs which convey, more often than not, a rather rigid vision of individual rights as determined once and for all in some respects at least? This is far from certain. Even though the ontological controversy has a direct influence on the choice of the analytic or philosophical apparatus, it does not necessarily convey a substantial view of the content or goals of rights. It is possible that the action-centred conception of rights has a special

appeal for thinkers with a libertarian bent or with a taste for natural-law theorising, but such associations are not necessary.

None the less, it remains true that such a view of rights seems well-suited for the non-manipulable, effective protection of rights, because rights are defined independently of social procedures. No matter the linguistic ambiguities in the allocation of rights, it puts limits on the potential changes in the detailed prerogatives of individuals, in some respects at least: they cannot be altered through strategic initiatives in political or judicial decision-making.

On the other hand, one of the distinct advantages of the social choice theoretic format is that it enables us to put all norms (legal and otherwise) on a par. Various codes of conduct or urbanity, from politeness and etiquette to full-fledged legality, can be described as properties of social choice procedures for decision-making. Then it is in principle conceivable, in some circumstances, that codes of conduct which are not legally binding (although they embody fairness, equity or decent restraint) could place limits on the exercise of legally conferred individual rights. This is not so obviously conceivable if rights-exercising is described as a game in which individuals are able to place *a priori* constraints on subsequent interactions, come what may. If we start from such a representation of rights, accommodation of the interests of others is hardly a collective issue. It can only be based on individual preferences and reciprocal exchanges of rights, motivated by mutual benefit.

If we remain faithful to the social choice format (convincingly advocated by Sen, 1992), we must recognise that the rules of social interaction, which stand out as procedural properties in the theory, partly derive their acceptability from the features and plausible evolutions of the concrete situations they regulate. The acceptability of Sen-type liberalism must be investigated in this perspective. Is Sen's minimal liberalism an acceptable necessary condition for freedom in certain social settings, but not all? Or should we accept Sen's liberalism conditions as necessary conditions which hold for *any* conception of rights, and are applicable for *any* social procedure which incorporates individual rights among other characteristics?

Sen's so-called "liberal" dilemma provides one reason to answer the last question by the negative. It shows that, should we accept minimal

liberalism, we should also accept the violation of the Pareto principle. Of course, we may infer from this result that the Pareto principle is not innocuous, as it clashes with an apparently mild requisite on rights. This was indeed Sen's original interpretation of the paradox. But we may also remember that the Pareto principle is a very mild consequentialist requisite on social choice procedures. Alternatively, we might be led to conclude that the paradox shows the highly problematic character of "minimal liberalism" when applied to all systems of preferences (as required by the "universal domain" postulate) and all interaction contexts.

b. The plausibility of "minimal liberalism": a contextual analysis

It is true that in some contexts, minimal liberalism seems innocuous. If regulation by means of rights-allocation occurs in a large society (where it is comparatively easy for other people to leave my behaviour unnoticed), minimal liberalism looks very innocuous indeed: I can see no point in letting some external power preventing me from choosing the book I want to read, or the colour of the walls of my bedroom (if I am the only user of the room). Substantial instantiations of Sen-type liberalism come to mind quite easily and they look equally compelling.

On the other hand, in a situation such as that described in Sen's Lady Chatterley example in 1970, it is hard to imagine that people could accept their being trapped in suboptimal results out of respect for absolute individual rights on the control of collective alternatives. Alternatives are collective, after all, so that "local" dictatorship (on only one pair of options) really amounts to arbitrary power. Sen's classic Lady Chatterley example goes like this: Prude and Lewd have a copy of Lawrence's famous novel and there must be a collective choice among alternatives a, b, c (Figure 1).

	L reads	L doesn't read
P reads	_	a
P doesn't read	b	c

Figure 1

Let us suppose that P prefers c to a and a to b; and L, a to b and b to c. Insofar as reading a book is a personal choice (with respect to which freedom of choice is warranted by liberal States), giving absolute control to P with respect to the choice between a and c, and giving absolute control to L with respect to b and c may seem natural as an application of Sen-type "liberalism" or "minimal liberalism". Indeed, the only difference between a and c, for example, is that in one alternative P reads the book while in the other one, he doesn't.

Then, socially speaking, c is preferred to a, and b is preferred to c. But the Pareto principle would require that a be preferred to b for society at large. Then there is no definite choice for this two-individual society, even though both individuals have well-defined, consistent preferences.

In this situation (which illustrates Sen's theorem on the impossibility of a Paretian liberal), let us take notice of a disturbing fact. Agent P is "willing to read" when a certain pair of social states is considered (he prefers staying at a or shifting to a, rather than going to b or staying at b), and "willing not to read" when another pair of social states is considered (he prefers staying at c or shifting to c, rather than shifting to a or staying at a). And similarly for agent L: comparing a and b, he prefers not reading to reading and comparing b and c, he prefers reading to not reading. Starting from a situation where he reads and L doesn't, P would readily shift to the social state in which nobody reads. On the other hand, he would rather stick to the status quo than shifting to a social state in which L reads and he doesn't. In such a situation (or interpretation of the formal social choice structure), it is by no means obvious what being "willing to read" means in the first place. Hence the social protection of the individual's ability to read if he wants to read has no precise meaning either.

In the Lady Chatterley example, is it certain that the social choice situation, as it has been specified, really makes unilateral, individual two-way decisiveness on one pair of options a compelling and minimalist requirement? This context and the ordinary context of interaction between people with no strong mutual link in a large society are quite different from each other. The main difference appears to be as follows. In the "large society" context, the external effects of individual action are somehow trivial, so that the unilateral determination of some

aspects of social states really appears innocuous. The changes it brings about in the world can be eschewed or nullified at low or zero cost by the other people. In settings of the other kind, by contrast, external effects are such that we cannot plausibly ignore them.

The meaning of rights and freedom should—and probably does—evolve in an adaptive way, in response to changes in the collective choice problems which are concretely encountered. Perhaps our schemes of interaction co-evolve with general solution schemes for arbitration of conflicting claims, in the adaptive processes of the species or subgroups. At any rate, a philosophical paradox like Sen's "liberal" dilemma may help us to correlate the interactive contexts and the concepts of rights: it reveals a logical impossibility in the conjunction of two claims, and one of the two claims has to do with protected freedom (or well-recognised rights). But the concept of rights which is outlined in the paradox may be relevant for certain contexts, not for others. In some cases, it could be criticised, arguing from the conflict with the unanimity (or Pareto-principle) requirement. Or alternatively, it could be experienced as defective in the face of controversial rights-exercising and this could be hypothesised to be a driving force behind normative changes in the allocation of personal prerogatives. Thus, reflecting on how people manage to live with rights in spite of the paradox might give us insights into the adaptation of our concepts of rights to varying circumstances of interaction.

It would be awkward to describe the choice of a history student to read *Mein Kampf* in a synagogue as a pure matter of unconditional rights-exercising—even if the choice situation is described in such a manner that the only choice left to an individual is to read or not to read the book. Reading a book of one's choice is clearly an individual right. But the alternative of reading or not reading *Mein Kampf* in a synagogue typically makes sense only in a social perspective: it is significant to abstain from it, out of respect for the preferences of others (and in the case at hand, out of respect for the well-justified moral expectations at the heart of their preferences), even if some kind of pressing historical curiosity suggests reading it during a religious ceremony.

Consider this other example: two persons are in a closed compartment in a train and, given that there is no available entertaining activity,

the only choice left is to smoke or not to smoke; this is reflected in the description of the social choice problem they face. In such a context, giving any individual an exclusive control on at least one aspect directly relevant to her (his) own person would amount to let him impose smoking upon the other. This would be highly problematic by any standard.

Presumably, different social contexts call for different interpretations of freedom and rights. In large communities with weak external effects, easy protection of privacy and prohibitive negotiation costs (in the sense of Buchanan and Tullock, 1962), freedom and rights usually revolve around unilateral capacities to fix various features of the world either prior to social choice procedures (as Nozick would have it) or within social choice (following Sen). On the other hand, in such a setting, it is hard to believe that external effects will give rise to the kind of virtual clash between the Pareto doctrine and rights Sen's dilemma displays. Although it is logically possible, it can be expected that it won't occur.

c. Types of social environment

In an open social environment, externality problems caused by differences of judgment about a book are unlikely to give rise to a dispute over rights which is motivated by efficiency concerns. It seems quite efficient for most purposes, on the face of it, that individuals should be left capable of decentralised initiatives on matters which are of direct interest to them and which can easily be ignored by the others (irrespective of the moral judgments of individuals on the actions of the others). Allocations of rights which embody traditional "negative-freedom" rights can be analysed as providing guarantees with minimal (easily justifiable) constraint and little loss for the constrained (Picavet, 1997; 2002; 2005; 2011).

The whole picture looks very different on the scale of small, integrated communities with strong links between the members. In such contexts, it is difficult to believe that people will leave local dictatorships unchallenged, if these prerogatives regularly prevent them from reaching efficient solutions. If an initial allocation of individual rights makes it possible for individuals to act out of their own preferences with potential adverse consequences for the others, it is to be expected that rights-exercising will be equated, in some cases, with inadequate

or deficient (not just impolite) decision-making. As insights from the sociology of normative dynamics would suggest, the initial allocation of rights will certainly be criticised and, if the power balance in society makes it possible, it will be changed.

In such contexts, even if we start from a pattern of unconditional individual rights to self-determination (with respect to "personal" matters) it is predictable that people will waive their rights to facilitate reciprocally beneficial collective decision-making, unless strong deontological commitments (such as a belief in the obligation to better one's life through narrowly personal choice) prevent them from reasoning in a consequentialist manner from a collective perspective (see the eloquent fishermen counter-example in Baron, 1998: 23–26).

The very meaning of freedom and rights can be modulated in response to various settings of interaction. Individual rights are usually described in general, hence flexible terms. If the problems at hand are the various collective choice exercises in a family for example (such as choosing a TV programme or rearranging the furniture), freedom can hardly be identified with the fact of having the last word. Freedom surely cannot consist in the capacity to impose a TV programme without paying attention to the reluctance of the other family members (no matter the precise content of the social choice to be made). The individual capacity to watch a specific TV programme becomes a pure matter of individual freedom only if there is some technical device which makes it possible for the other members of family not to be disturbed, or hindered in their own choice, by this specific individual choice.

Preference-based individual choice need not imply the disregard of social features outside one's own perception of the differences between objects in a pair of possible social states. An individual may prefer x to y without insisting that x should materialise. For example, it is possible that several individuals agree to materialise a third alternative, z, which they prefer to both x and y. And it is possible that some reasons, grounded in fairness standards, prevent an individual from insisting on his preferred choice (acting otherwise would amount to inadequate decision-making). Then all the relevant social reasons for choice should be taken into account and rights-exercising should be appraised from a moral point of view or, more generally, from a normative point of

view—even though there might exist convincing general (non-contextual) reasons in favour of the considered rights. This line of reasoning does not turn into an impossibility the fact that *some* individual rights might appear truly related to personal domains, and worthy of social protection in this respect; but it supports the conclusion that contextual characteristics are relevant when we try to reach an assessment of the correct or incorrect nature of decision-making, in cases where the decisions taken truly qualify as cases of rights-exercising.

The ability to have the last word on some pairs of alternatives doesn't appear to be a necessary condition, unless we consider those choice problems which turn out *not* to be of significance to more than one person (such as sleeping on my back or on my belly), and hence, not susceptible to external effects and hardly a suitable subject matter for explicit social choice procedures. With respect to small-world situations, freedom will typically consist in procedural claims, especially, in the fact of having one's word to say in the decision-making process, and withdrawal rights (opportunities not to interact any more). It will typically never consist in getting what you want no matter the opinion of the others. Here a question naturally arises: if Sen-type liberalism is such a counterintuitive condition in small groups with strong external effects inside the group, how comes that we are sometimes willing to describe it as a condition that is "minimal" indeed, no matter the precise context of interaction?

Part of the answer, I believe, has to do with the fact that when we think about minimally specified liberalism, we typically apply it first of all to such matters as sleeping on one's back or one's belly or shifting from one thought to another—matters which are really not of interest to (or distinguishable in the eyes of) more than one person. Such matters may account for the differentiation of options in the standard (arrovian) social choice setting. But insofar as matters such as these are under scrutiny, the Nozick-Gärdenfors argument (about rights as ways of determining features of the world prior to any collective choice) has maximal force. Why not let people decide on such matters beforehand, prior to any kind of collective interaction?

The normative ambiguity involved in the paradox is clearly felt in Sen's own discussion of the Lady Chatterley example. Freedom is

supposed to be important in the form of one's being capable to do or not to do something which is of interest to oneself (as in the classical approach to individual rights in terms of *capacitas* since the time of Grotius and Leibniz). That's why it seems minimal: it captures the ability to act as you like in at least a very tiny personal sphere, with a view to a desirable subset of the possible results. None the less, as the example suggests, the "personal sphere" is not so personal after all. Given the externalities in the situation, the so-called personal spheres are integrated to a considerable extent, so that exclusive personal control on any such sphere appears deeply problematic upon examination. In contexts of this sort, the externalities can also be viewed as causal factors in the emergence of formal or informal social norms, as explained by sociological theories of the dynamics of norms (Opp, 2001).

3. INDIVIDUAL RIGHTS AND PATTERNS OF SOCIAL INTERDEPENDENCE

a. Reinterpreting the "liberal paradox"

Domain restrictions provide classical escapes from Sen's so-called "liberal paradox" (Pétron-Brunel, 1998, Igersheim, 2004). Usually they are formulated on individual preferences. The same pieces of criticism regularly appear: why suppose some people have this or that kind of preferences? And if the restriction is motivated by moral reasons, what is the point of supposing that morality as the theorist thinks of it is better than the morality of agents who happen not to be in the restricted, acceptable domain?

Domain restrictions are perhaps more acceptable if they are based on social mechanisms or types of interdependence which are at work in the case of certain rights. If we are able to identify some mechanisms, we can examine their implications for the mutual dependence of the agents. And if we characterise the association of such mechanisms with a set of rights as *efficient*, then we have a reason to say that the recognised individual rights have co-evolved in an efficient way with the kind of social life in which the identified mechanisms are effective.

Let us suppose that human groups have evolving concepts of rights, which are partly developed in response to coordination problems, in the

sense that some norm of efficient behavioural coordination places constraints on the selection of concepts of rights. Then it is quite reasonable to suppose that individuals who encounter situations such as "Lady Chatterley" will not be satisfied with a "minimal liberalism" constraint on their concepts of rights. Rather, they will develop alternative concepts—perhaps participatory or negotiation rights (rights at the procedural level), or rights to a fair (or mutually advantageous) solution. The real but limited flexibility of the concept of a "right" makes room for such adaptation. Two facts are worth noticing in this respect:

— (i) some kind of minimal individual control always escapes efficiency problems of the kind revealed by the Liberal Paradox because, even in very small and integrated societies, some initiatives are put under individual control, which involve no barrier to any other member's promotion of her (his) own welfare through individual action—sometimes because the actions are not even distinguishable in the eyes of other people. A typical example is Sen's example of the right to sleep on one's back or one's belly. In this respect, "minimal liberalism" in a social-choice framework is just a reformulation of Kelsen's classical argument to the effect that no system of constraint whatever can block every kind of personal initiative in those areas which are not explicitly an object of constraint in the normative constraint order (Kelsen, 1960: sec. I.6.b.ε).

  This somehow trivialises the defence of various rights' concepts through a test of compatibility with minimal individual control. If the attractiveness of "minimal liberalism" really comes from this kind of ultra-minimalist requirement, then many rights-restricting efficiency recipes should appear acceptable from the point of view of such libertarian "minimalism". Sen's paradox warns us that it is in some sense problematic to push libertarianism a step further.

— (ii) many rights give guarantees for the results of personal initiatives which do involve externalities, in the sense that both the actions and the associated guarantees place binding constraints on the other agents' ability to promote their well-being through their own initiative. But social mechanisms are available, through

which individuals can limit the harm they inflict to one another, thanks to contingent aspects of typical interaction in large groups (such as the easiness of being unaware of what other people do in some regions of their lives, and the weak dependence of the results of one's actions on the precise actions of specific persons). These mechanisms can be reinforced by a liberal culture of equity, toleration and self-restraint (well-captured in a formula such as "this is none of my business", or "I should pay attention to the needs or values of others", or else, the so-called "golden rule" of ethics), not only by the voluntary exchange of privileges along the lines of Barry (1986).

Under certain conditions of interaction, then, we can say that rights-exercising in given sectors of an actor's life revolves around rights with the following property: these rights allow the unilateral promotion of each bearer's well-being while affecting in no significant way the comparative attractiveness of the actions available for the other actors. This property is well connected with a classical, highly-regarded view about freedom. It warrants that there is a definite meaning, for each individual, in the willingness to exercise a right or not (with the aim of making the transition from one social state to another effective). This meaning is definite because it is independent of the details of social interaction. In other words, we are able to make sense of the notion that freedom is "freedom to act", or the ability to determine features of the world in an intentional (or voluntary, purposeful) manner, and in a way that effectively promotes one's good (or one's welfare or preference-satisfaction) without conditioning this ability by the details of a social interaction scenario (one key aspect of the Nozick-Gärdenfors critique). Unless this is the case, it is not obvious that individuals should be given an absolute, unilateral control of their own action on the basis of an independent right to do this or that according to their own personal preferences; this would be quite strange indeed, as the "Lady Chatterley" example aptly demonstrates, in situations where such a right has no clear meaning, because the attractiveness of certain actions is essentially dependent upon the choices (and indirectly the preferences) of other agents.

The discussed property connects up with the ordinary notion of freedom. Its proposed interpretation relies on mechanisms that seem to be present in the way we deal with externalities in large, rights-based political societies. And finally, it blocks a mechanism that is present in the canonical example of the Liberal Paradox—a mechanism of mutual dependence which endangers the connection between rights-exercising and "freedom" in an ordinary sense.

b. Requirements for autonomous choices, preferences and interdependence

I suggest we take as rough guidelines the following features that minimally specified individual controls should have in social choice in any event:

(i) having such a control enables an individual to determine some features of social choices among alternatives, in a way which cannot be made irrelevant by similar capacities of other individuals;
(ii) the choice to exercise such a control in social choice should be valuable for the person, irrespective of the choices by other persons in equivalent minimally specified spheres of personal control in social choice.

Roughly speaking, *(i) may be considered a requirement of non-domination, or social respect for autonomous choice*: the individual freedom granted by means of individual rights in personal domains should not be screened off or reversed in social choice by equivalently personal rights-exercising on the part of other individuals. Such a statement can be made more precise by a closer examination of the "problematic" cases behind Sen's paradox. This is in line with Risse's recommendation (2001: 183) that we focus on the circumstances in which conflicts of principles occur.

Of course, I may accept that some of my private-sphere choices be made irrelevant by some choices of other individuals, especially those choices of other people which bring about notable changes in society at large. For example, I may wish to control the colour of the kitchen of a poor-quality apartment I live in, without opposing a public

decision to replace the old building with a new one, in a scheme which includes preferential rents for old-time district dwellers, in such a way that my preferences about the colour of the kitchen in the old apartment becomes irrelevant to me. But the same will not hold with respect to private-sphere decisions of other people.

Surely it is a matter of principle that my decisions in a private sphere, if they are recognised as such in social choice rules, should not be made irrelevant by *equally private* decisions of other people. Otherwise my own private sphere, as it is recognised in public life, would be publicly considered subordinate to the private sphere of other people in some cases (and this kind of subordination could be reciprocal in some cases).

Condition (ii) is mainly to be viewed as a test of minimal interpretative relevance. If it is satisfied, then minimal rights in private spheres are associated with actions which minimally matter to individuals in the following sense: they are associated with choices which cannot be made valueless by similar private-sphere actions or choices on the part of other individuals. This means that the social-choice problem at hand is considered with a view to those social situations in which it makes sense for individuals to say that they "want to exercise a right", irrespective of the details of the actions of other individuals in private spheres. Concerns for "minimal" rights are typically associated with minimal freedom of action (or margins for autonomous action), and minimal freedom of action has to do with being able to do whatever one wants to do in a given action domain. In order for this to make sense, we must check that the situation at hand involves a sufficient degree of individual, independent motivation for the choice of actions. This is expressed by condition (ii).

My intention is not to suggest that individual preferences should undergo some kind of reform or censorship so that the "problematic" patterns do not come up. Rather, condition (ii) is meant to provide guidance for the analyst, when she looks at real-world situations. Not all "minimal" normative concepts will do for any situation.

In situations such as Sen-type "problematic" cases, both (i) and (ii) are violated. Preferences are perhaps not to be blamed in the first place. Quite simply, we are in situations where Sen-type individual controls will not be expressive of minimal social respect for individual choice

in restrictively specified domains. Perhaps they can express something else, but they can't express just this, and we should be concerned precisely with this fact insofar as we deal with minimally specified individual rights.

c. The dependence of rights on patterns of social interdependence

The social allocation of rights and freedoms is typical of human coordination as it is at a given stage of the biological evolution or adaptation of the species. Yet our human languages are often ambiguous on essential matters, and freedom, rights-based and liberty ideals are no exceptions. It is probable that our linguistic or conceptual conventions about such values and norms evolve under the pressure of coordination dilemmas, especially efficiency problems. In his bold naturalistic attempts to explain the social emergence of privacy norms and privacy domains, Jack Hirschleifer (1979) observed that language is essential to social coordination in the human species. Linguistic variations and conceptual uncertainties are widely recognised as major source of disagreement about rights and freedom as well as the allocation of private domains of unlimited freedom of choice.

Remarks such as these suggest a fresh look at domain restrictions which may help understand the practical relevance of rights in spite of the conflict of rights with efficiency (classically encapsulated in Sen's "liberal paradox"). Perhaps we should face more squarely the fact of the essential variation of our concepts of freedom and rights according to the contexts of interaction. This is a motivation to investigate the social mechanisms through which the practical importance of externalities of the sort involved in Sen's paradox is kept to reasonable proportions.

Indeed, it is widely acknowledged now in economics, in the continuity of pioneering investigations by Coase (1988) and others, that the allocation and the social definition of rights (notably property rights) can evolve in response to efficiency problems. In the social world, however, property rights are not just technical allocations of access and control.[2] Our social concepts of rights and freedoms are supported by normative judgements about freedom and liberty (and if we insist that they are supported by interest only, we should acknowledge that this interest is partly constituted by the respect showed, and protection

granted, to activities, states of the person or forms of life which are considered legitimate). Hence it may be argued that our accepted explanations of the evolution of rights in response to efficiency problems are still incomplete as long as we lack an explanation of the adaptation of our very notions of freedom (liberty) and rights in the processes of social life. The social sciences have delivered a lot of insightful remarks on the relationships between efficiency and freedom. The "liberal paradox" is a notable and celebrated example. We know and conjecture many things about the relationships between efficiency and rights. But it must be allowed that there is still little integration of the best conjectures about the adaptive aspects of rights (and obligations, and responsibilities) on the one hand, and our normative discussion pertaining to the importance of rights on the other hand.

Some authors advocate a view of human agency which is predetermined by social obligations. Taking this perspective, of course, the correlativity of the exercising of rights with social norms which give it its value is quite obvious. But this essentially correct conclusion should not be bought at too high a price. Ballet *et al.* (2007) write:

> … In real life, especially in traditional societies, freedom often depends on first satisfying one's own personal obligations towards others. It is this capability of first fulfilling one's obligations that subsequently generates the capability of choosing what one wants rather than *vice versa*.

This is perhaps difficult to accept. The fact that the fulfilling of one's obligations is a precondition for getting good results through the choice of certain actions does not imply that the capability for preferential choice is in any sense "generated" by social norms and normative relationships among the individuals. One of the powerful warnings of sociological reasoning against this implication is to be found in the Weberian analyses of the way individuals, in full awareness of the connection between social institutions and their own interests, are able to influence the predominant interpretations of norms and the operation of institutions (Cherkaoui, 2006: ch. 6, sec. C). The social structure of responsibilities is thus influenced by individual initiatives.

4. CONCLUSION

It is widely thought that the preferential choice of an individual in various action domains should benefit from a degree of social protection, but individuals can be led to renounce choices of this kind for good reasons in the light of collective properties, such as shared benefits, fairness, polite attitude and so on. For this reason, the individual right to influence social outcomes through preferential choice can hardly be considered an *a priori* binding constraint for morally admissible social choices; other values matter as well for the purposes of normative evaluation (if not for the purposes of the assessment of legality, whenever the considered rights are couched in the law). Accordingly, it is perfectly advisable to examine the normative or moral value of individual choices with respect to such norms as fairness, shared advantage, etc. even though they qualify as rights-exercising instances in the first place. Thus it seems that respect for private choice, and the institutionalisation thereof, do not commit us to a view of individual rights as unconditionally binding constraints for all contexts of interaction. Rights-exercising is a relevant object of normative evaluation and the complexity of the exercise cannot be simplified too much.

NOTES

1. Support of the Agence Nationale de la Recherche (projects DELICOM, JC-JC05 and 3LB, NT-05) is gratefully acknowledged. Both projects have benefited from the environment of research team "Philosophies contemporaines" (E.A. 3562), especially in the NoSoPhi group, at Panthéon-Sorbonne University. I also thank the Philosophy and Political Science departments at Oslo university (Norway) for their hospitality on the occasion of a colloquium on the issues of fairness (reflecting the interests of the "Mind and nature" research project) where I had an opportunity to present an earlier version of this contribution.
2. To be sure, there are unification gains on the descriptive level if we describe them just in this way, as exemplified by Coleman (1990).

REFERENCES

Arrow, K.J. (1951) *Social Choice and Individual Values*. 2nd edn., 1963. New York: Wiley.

Ballet, J., Dubois, J.-L. and Mahieu, F.-R. (2007) "Responsibility for each other's freedom: Agency as the source of collective capability", *Journal of Human Development*, 8(2): 185–201.

Baron, J. (1998) *Judgment Misguided*. New York and Oxford: Oxford University Press.

Barry, B. (1986) "Lady Chatterley's lover and Doctor Fischer's bomb party: Liberalism, Pareto optimality, and the problem of objectionable preferences", in: Elster, J. and Hylland, A., (eds), *Foundations of Social Choice Theory*. Paris: Editions de la Maison des Sciences de l'Homme and Cambridge: Cambridge University Press.

Brunel, A. and Salles, M. (1998) "Interpretative, semantic and formal difficulties of the social choice rule approach to rights", in Laslier, J.-F., Fleurbaey, M., Gravel, N. and Trannoy, A., (eds), *Freedom in Economics. New Perspectives in Normative Analysis*. London: Routledge, pp. 101–111.

Buchanan, J. and Tullock, G. (1962) *The Calculus of Consent*. Ann Harbor: The University of Michigan Press.

Cherkaoui , M. (2007) *Le paradoxe des consequences*. Geneva and Paris: Librairie Droz. Published in English as *Good Intentions: The Paradox of Unintended Consequences*. Oxford: Bardwell Press, 2007.

Coase, R.H. (1988) *The Firm, the Market and the Law*. Chicago: University of Chicago Press.

Coleman, J. (1990) *Foundations of Social Theory*. Cambridge, MA: Harvard University Press.

Gärdenfors, P. (1981) "Rights, games and social choice", *Noûs*, 15: 341–356.

Gaertner, W., Pattanaik, P. and Suzumura, K. (1992) "Individual rights revisited", *Economica*, 59: 161–178.

Hirschleifer, J. (1979) "Privacy: Its origin, function and future", *Working Paper series*, No. 166, University of California at Los Angeles, Department of Economics.

Igersheim, H. (2004) *Liberté et choix social. Contribution à l'analyse de la liberté en économie normative*. Doctoral dissertation, Université Louis Pasteur, Strasbourg, France.

Kelsen, H. (1960) *Reine Rechtslehre*, 2d edn. Vienna: Franz Deuticke (1st edn. 1934).

Nozick, R. (1974) *Anarchy, State, and Utopia*. Oxford: Basil Blackwell.

Opp, K.-D. (2001) "How do norms emerge? An outline of a theory", in Boudon, R., Demeulenaere, P., Viale, R. (eds), *L'explication des normes sociales*. Paris : Presses Universitaires de France.

Pareto, V. (1905) "L'individuel et le social", in Claparède, E. (ed.), *Congrès International de Philosophie (Genève, 1904)*. Geneva: Kündig. Also in Pareto, V., *Mythes et Idéologies*, Busino, G. (ed.). Geneva: Droz, 1966.

Pétron-Brunel, A. (1998) *Contribution à l'analyse des droits dans la théorie des choix collectifs*. Doctoral dissertation, Université de Caen, France.

Picavet, E. (1997) "Rational agreement on human rights in a pluralistic setting", *Rechtsphilosophische Hefte*, 7: 109–135.

——— (2002) "Conflictos de intenciones y racionalidad de los sistemas de normas" (tr. by Bernardo Bolanos-Guerra), *Isonomia*, 16(April): 113–147.
——— (2005) "L'intervention du raisonnement mathématique en philosophie politique". *Mathématiques et sciences humaines/Mathematics and Social Sciences*, 43(4): 43–65.
——— (2011) *La Revendication des droits*. Paris: Classiques Garnier.
Risse, M. (2001 "What to make of the liberal paradox", *Theory and Decision*, 50(2): 169–196.
Sen, A.K. (1970a) *Collective Choice and Social Welfare*, chaps. 6 and 6*.
——— (1970b) "The Impossibility of a Paretian liberal", *Journal of Political Economy*, 78: 152–157.
——— (1992) "Minimal liberty". *Economica*, 59: 139–159.
Sugden, R. (1985) "Liberty, preference, and choice". *Economics and Philosophy*, 1: 213–229.

17

FUNCTIONALISM, FIELD THEORIES, AND UNINTENDED CONSEQUENCES

Stephen P. Turner
Department of Philosophy, University of South Florida

One of the virtues of Mohamed Cherkaoui's *Good Intentions* (2007) is that it separates the problem of unintended consequences from functionalism, and shows its power as an autonomous concept. In what follows I will take up one key idea in the book, the limitations of instrumental rationality. This is closely connected to the notion of unintended consequences: acts with unintended consequences are often acts where, because of our limited knowledge, the consequences are not known in advance nor expected because of our limited knowledge. Cherkaoui identifies different forms of the unintended social consequences of action. The effect of this is to show that functionalist arguments are unnecessary: that a better account of unintended consequences, completely casual in character, suffices to account for a wide range of phenomenon formerly thought to require, or be congenial to, functionalist explanation. This represents an important advance. Nor is its relevance restricted to the past, before functionalism became the love that dared not speak its name. It applies to the later, submerged forms of functionalism that attempted to evade the criticisms of functionalism.

Despite the very different trappings, Bourdieu never went beyond functionalism, as Elster has recently pointed out (2010). The functionalist pattern is one of reasoning in which the *consequences* of an action or pattern of actions explain the maintenance of the actions or pattern of actions. Often this requires theoretical invention, such as reasoning backwards from the existence of some social fact to a hypothesized sustaining force, which is then theorized as a kind of goal-seeking or end-producing agent-like force. This pattern of reasoning applies, evasions notwithstanding, to a wide range of social theories, including so-called field theories and Merton's "structuralism." But evasion is characteristic in these accounts. The evasions in the case of functionalism are failures to fully explain how it is that the consequences explain the cause. The evasions typically take the form of truncations or omissions, in which some crucial explanatory element is left obscure, left out, or left incomplete. As we will see, this is an important feature of "field theories." The evasions serve a purpose, however: they allow the theorist to avoid appealing to problematic theoretical inventions of hidden forces.

LEFT FUNCTIONALISM AND THE MYSTERIES OF TELEOLOGY

Traditional functionalism validated apparently irrational practices or institutions by showing that they had unrecognized beneficial effects, usually for "society" or for people other than those who were engaged in the activities of the institution itself. Left functionalism took society to be oppressive, and thus interpreted practices and institutions as secret oppressors. It was not claimed that the practices were intentionally or overtly oppressive, but that they were oppressive in their consequences, or in their hidden essence. Claiming they were oppressive in their consequences is not a claim that that fact explains the pattern. The claim that they were oppressive in their hidden essence, however, does: it implies something like a true purpose and force. This kind of reasoning characteristically took a theoretical turn, in which the oppressive agency was theorized as being itself a concealed cause that could be named and described, such as *habitus*. In this respect Left functionalism

resembles conspiracy theory. One can invent explanations of something one condemns, and then attribute it to this malign force. Nothing is happenstance. Any large effect must have a large cause. The cause is intrinsically hidden. For functionalism in its Bourdieusian form, these secret forces operate within and through ordinary agents–intentionally acting people. Yet the forces are not blind, because they adjust to continue to produce the result that is being explained. But the intentional agents whose actions are the proximate producers of the result are blind to these forces. Teachers, for example, think they are doing something else– enforcing standards, educating, and so on–when they are instead unwittingly enacting the inscription on the bodies of their charges of a *habitus* which ensures the continuation of class distinction.

Habitus is a characteristic theoretical object in this tradition. There is a banal and uncontroversial form of the concept, that appears in Weber, and means merely the collection of habits of an individual. It has no collective causes, though it might have effects on others and on collective outcomes. Habits are acquired in the normal manner: through experience and such things as actions or forms of reasoning becoming second nature, habitualized. For Bourdieu, however, *habitus* is something more. Not only is it the source of individual dispositions, these dispositions serve social or collective ends, and, crucially, are part of a process that reproduces oppressive effects. A *habitus* is a kind of collective agent that expresses itself in individual dispositions produced by the socially generated individual *habitus*.

One can see here the beginnings of the issue. The fact that there are things that people (in this case teachers) do that have unintended consequences is part of this story. But to get the teleological directionality, the continuing adjustment to produce the same unintended results, which is essential to functionalist explanation, there must be more. The functionalism is concealed in the term reproduction. What is being "reproduced" are the effects, and they are reproduced through the means of the theoretical object, the *habitus*, which is connected to the effects not merely causally, but in a way that sustains the effects. In short, for reproduction in the relevant sense, i.e., not merely copying but changing to continue to produce the relevant effects, there must be feedback. But the feedback cannot be through the medium of the adjustment of action to

conform with individual intentions, in the manner of individual teleology, the teleology of purposive action. The intentions and dispositions themselves are expressed individually but controlled through some sort of collective process.

The difficulty with these explanations and their variants is in making sense of this collective process. Old fashioned functionalists "solved" this problem in a variety of ways. One is to hypothesize steering mechanisms of some sort in a dynamic system, and to infer the existence of these mechanisms from the fact that "society" manages to make adjustments, for example, to its environment or internal tensions. Needless to add, these were all highly abstract rather than concrete phenomenon. But there is a special difficulty with Left functionalism: the mechanisms are not only hidden, they must be systematically concealed from consciousness in order to work. Oppression and its analogues cannot flourish if they are recognized as such. But oppression works through intentional actions. What is required for people to perform the oppressive acts is a specific kind of false consciousness: misrecognition, which is the reason the causes are hidden.

MISRECOGNIZED *VS*. UNINTENDED

Despite their surface similarity, the concept of misrecognition and the concept of unintended consequences are very different in their implications. A side-effect of a drug given for a specific purpose is an unintended consequence. The ethical doctrine of double-effect is concerned with the distinction between intended and unintended consequences. In these cases there is a clear separation between the subjective fact of intent and the objective facts about consequences. In the case of misrecognition, some notion of essence or essential connection is a part of the concept. And this is where some sort of supra-intentional teleology is built in to the usage.

Weber's thesis in *The Protestant Ethic and the Spirit of Capitalism* is the classic example of unintended consequences. The Protestants in question were aiming for salvation. In doing so they became disciplined and consequently rich. Getting rich was never intended. But the casual connection was eventually recognized by the people involved. There

is nothing in the way of supra-intentional teleology in this account. There is no essence. There is no secret force inducing the religious beliefs as a means of producing the end of wealth. And although the consequences were eventually recognized, there is no feedback loop in which this recognition serves to sustain belief. On the contrary, there is a concern that wealth undermines belief. To be sure, there might be other effects that work in the other way: a person esteemed initially for being rich might be imitated and in the course of imitation someone might come to believe as a side-effect of imitation. All of the feedback, however, is on the conscious or "recognition" level. There is no secret path. People are conscious of what they are doing, even if they are oblivious to the side-effects. It is not essential to the causal story that these effects be hidden.

In the case of misrecognition, it is essential that the effects be hidden, or at least incidental to the main point of the activity, and thus hidden from the attention of the agent. The thought is that the people performing the relevant acts would not do so if their real significance was manifest. If teachers understood that by enforcing rules of grammar they were actually (and essentially) reinforcing class distinctions, they would rebel. So something needs to do the hiding of these effects. And this something is ideology. But merely confusing people about effects–making them unintended, so to speak–is not enough. For ideology to do its work, it must hide: overcome the reality of the effects that normally would be apparent. Mere unintendedness is not enough. A kind of active participation in the process of self-delusion is necessary. And the process needs to be sustained, and even to adjust to new circumstances in order to continue hiding effects. Error is not enough either. Error can be overcome, either advertently or inadvertently. Error is a kind of natural accident: intelligible, but recognizable (under the right circumstances, such as the availability of new data) by the agent as wrong. It may be systematic, for example as a result of cognitive biases. Or it may be propagated intentionally, by believers or by people who know it is error. But error is not self-sustaining.

Bourdieu's account of *doxa* and *habitus* trades on the notion of systematic error. The world of experience given to us as a result of our shared *habitus* biases us toward seeing what is in fact the product of

history as "natural" and this systematic error is at the foundation of our beliefs, which we are systematically in error about because we do not consider the conditions of the possibility of belief. We confuse second nature with first nature, and as a result of the phenomenological immediacy of experiences conditioned by second nature, we ignore the conditioning itself. In Lacan, the source of the term "misrecognition," we get a straightforwardly teleological explanation complete with feedback mechanisms: misrecognition is a systematic means by which the ego protects itself from shattering in the face of the arbitrariness of the world. The work of protection is unconscious, the mechanisms of feedback and revision is thus hidden from the agent, and necessarily so: conscious self-deception would no longer be effective self-deception.

Bourdieu needs a collective analog to this teleology and something to substitute for these unconscious feedback mechanisms, and indeed he asserts exactly this in his analysis of Mauss on gift exchange. "It is clear", Bourdieu insists, "that individual self deception [which characterizes each party's position in the logic of exchange, such as the gift exchange described earlier] is only possible because it is supported by a collective self-deception … This collective self-deception, furthermore, "is only possible because the *repression* from which it arises … is inscribed, as an *illusio*, at the foundation" of exchange. The entire economy of exchange is based "on a collective labour devoted to maintaining misrecognition with a view to perpetuating a collective faith in the value of the universal, which is simply a form of individual and collective bad faith (in the Sartrian sense of lying to oneself)" ([1997] 2000: 192). This collective labour is entirely metaphorical, of course: if there were real "labour" it would be conscious and intentional, and this would defeat the purpose of the misrecognition argument because it would not be hidden. But what the phrasing also conceals is the lack of any teleological force operating at this subintentional level that is analogous to the Lacanian ego protecting itself. The idea that there is an *illusio* at the foundation of the form of exchange is about essences: the essence of the form of exchange is other than the manifest features of the form. The essence is captured by Bourdieu's description.

THE PROBLEM OF AMBIGUITY

Because functionalism fell so far out of favour, and the perspectives that replaced it appeared on the surface to be so different, it would be useful to clarify some of the issues with these explanations, issues that carry over to Left functionalism and "field theory." The issues are apparent with Merton himself, and openly discussed in the Merton literature. Jon Elster pointed out one of them in these terms:

> Merton never states in so many words whether the task of functional analysis is to *explain* social phenomena or, more modestly, to identify and describe phenomena that might otherwise be overlooked. Clearly, manifest functions—intended and recognized consequences of action—have explanatory power. The question … is whether latent functions—the unintended and unrecognized effects of action—can explain the actions of which they are the consequences. (1990: 130)

This is indeed the question, for Bourdieu and others as well as Merton. The issues appear at the heart of Merton's most famous arguments, such as his analysis of machine politics. As Elster points out, Merton

> … refers to 'consequences [of a structure] which may provide basic social support for the structure.' With respect to the political machine he claims that 'the functional deficiencies of the official structure generate an alternative (unofficial) structure to fulfil existing needs somewhat more effectively.' The first statement suggests that consequences are important in maintaining (and thus explaining) institutions, the second that they are instrumental in creating (and thus explaining) them. (1990: 131)

Elster frames this as a problem of Merton's intentions: were they merely to describe consequences, or to use the consequences to explain the institution?

The larger issue is whether functional locutions and especially claims about "maintaining" can be employed without a commitment

to a view of society, or some other unit, having ends. The issue is not settled by quotations like the following, but they do allow for some clarification. Writing about the professional association of nurses, Merton comments that

> But one of the association's principal functions for the society of which it is a part is far from evident; it has often gone wholly unnoticed. This is its function, as one of the great intermediate associations, to help prevent the atomization of society into a sand-heap of individuals, each intent on pursuing his own private interests. Such an atomized condition is a step toward totalitarianism, which consolidates power over these socially disconnected individuals into a single center. (1958: 53)

This is a hidden effect that has to be pointed out by the sociologist, an unintended consequence, and a latent function. But it is not a case of misrecognition in the sense that nurses would recoil in horror and abandon the practice if these consequences were revealed to them. Nor is it clear that the consequence has much to do with maintaining the association, and Merton never hints that it does.

This, then, is a case of functionalist language used to simply mean unintended consequence. The consequence is collective, so it is not the kind of consequence that the people performing the actions would have had in mind, or perhaps could have had in mind. But it does not explain the creation or maintenance of the association of nurses. Is the case of machine politics like this: that it is the case that it fulfils some needs more efficiently, but that this is merely a descriptive fact about machines, not a part of the explanation of them? This is an important point: it is simply the case that the kinds of institutional arrangements that figure in functionalist explanations do not go on forever. Political machines of the kind Merton described have long since vanished. Did something change in some sort of unspecified feedback mechanism? Or was it instead that the conditions that produced the machines ceased? Is the "efficiency" of the machine a descriptive fact about the machine, or part of an explanation of why it persisted or won out against rival institutional possibilities?

THE CASE OF "FIELD THEORIES"[1]

Merton's model of political machines and Bourdieu's model of teachers contain omissions or problematic entities. In the case of Merton, it appears that he wishes to claim that the efficiency of the machine in delivering benefits contributes to its maintenance. But to complete this claim he would need to supply the relevant link: a feedback or selection mechanism, for example. This is nowhere to be found. All we get is a claim about benefits. In the case of Bourdieu, we get an appeal to a problematic theoretical object and process together with an implicit claim about essences related to misrecognition. The problem of omissions arises in particularly acute form in so-called "field theories." These are claimed to avoid the problems of functionalism. They do so, however, by giving up on the causal explanations and mechanisms that functionalism relies on. But is it really possible to have something explanatory rather than descriptive and avoid cause and mechanism? In what follows I will consider John Levi Martin's revealing exposition and defence of these claims.

The basic idea of field theory is that one can "explain" in a non-causal or at least non-mechanistic way by describing the features of a "field." The puzzle, however, is how a field explains. Martin suggests an analogy to the idea of magnetic fields. In this kind of explanation, it is not the properties of the objects in the field, but the properties of the field itself that do the explanatory work. Similarly, it is claimed, for field theories. The forces that do the explaining are in some sense in the field or part of the description of the field. But there is a familiar ambiguity here. Are we merely describing a world of existing forces, so that field theories are models which apply when they apply? If so, what are the conditions for their application, the conditions that make them fully explanatory? Are the forces in question enough to complete the explanation, or is something else needed? And are these "forces" problematic theoretical items, which are attributed unusual properties, but exist only to fill holes in the explanation?

Martin notes a "danger of … tautology" (2003: 8): "since fields are only known by their effects … it is tempting to proliferate invisible forces that 'explain' whatever it is that we otherwise cannot explain" (2003: 8).

This is of course familiar territory: Bourdieu and Foucault, not to mention Parsons, were fond of naming such forces or inventing theoretical objects with unusual powers to explain the phenomenon they attempted to explain. But these forces and theoretical objects, even if they were unfamiliar, explained in more or less familiar ways: they worked with mechanical analogies, or traded on the idea of presupposition.[2]

Field theories are more problematic. Martin uses Gestalt theory as a model close enough to social science to shed light on how field theory in social science might be understood. He argues correctly that Gestalt theory does not provide mechanisms. Indeed, this is one of the major criticisms of it. Nor are field theories either generalizations or generalizable. We do not know why they apply, or why they fail to apply. They differ from functionalism in that they avoid the characteristic nested structure of teleological explanation, in which ends become means to ends in higher or larger functional units. They are resolutely "local" or "concrete."

> the fact that the field at some place and time can be determined to be of a certain nature in no way implies that it must be this way—indeed, field theory, by never making explanation reach outside the field, must forswear any legitimating arguments that there is a reason why the field must be as it is. For this reason, field analysis is quite different from systems analysis, which, though it may stress the self-organization of the system, requires that the system be understood in contradistinction to an environment. But in field theory, explanation stops at the constitution of the field. (Martin 2003: 12).

The language explaining the notion of what it means for the model to apply is archaic: "determined to be of a certain nature." This is the language of essences. But the point is plain. Field theories are "applies when it applies" models. They forswear answering the question "why does it apply?"

But field theories, by renouncing any interest in causal mechanisms (however misleading this renunciation turns out to be[3]), find themselves on the horns of a different dilemma: either they are descriptive rather

than explanatory, or they are explanatory but explanatory in a teleological way–precisely as functionalism is. The issue can be seen clearly in the example of Gestalt psychology, which Martin repeatedly turns to as a model. Martin tries to navigate this dilemma as follows. He argues that these theories differ from normal functionalist sociology because the fields are local, the analysis is concrete, and the explanations do not depend on some sort of general societal end state. Is this possible? Not for functional explanation. To explain a given fact functionally, one must say what functional contribution it makes. The ends, the ends that do the explaining, may be local. But local ends must themselves be explained. The explanation might be mechanical or causal: a machine can be designed to equilibrate, as a thermostat does. If one has a complete description of the feedback mechanisms and so forth, one has a causal description of the device, and the preferred state is no longer the explanation but rather a feature of the causal processes themselves–a mechanical fact about the machine setting. But this kind of account converts a teleological explanation into a causal one.

Teleological explanations, explanations invoking ends, produce a hierarchy or sequence of ends, in which a local end is a means to another end. If we are to explain why animals have particular capacities by saying that it conduces to their survival, we invoke a preferred state, survival. But we then need to ask why survival is a preferred state. And we get some explanation like this: survival enables the genes to be transmitted. Genes being transmitted becomes the preferred state. The explanation is teleological. In the social case, there is always the question of whose ends are doing the explaining: whether they are individual ends (the interests of the voters who support the machine) or societal ends (such as integrating these voters into larger political order to produce stability, or some such explanation).

Field theories are neither casual nor teleological explanations in this sense. The question then becomes whether they are explanations at all. In the case of magnetic fields, the answer would be that the field description is incomplete: one needs to account for the forces themselves. A similar problem arises for Gestalt theory. Critics have complained that Gestalt theory is purely descriptive, inasmuch as it describes a perceptual state of Good Gestalt, but does not provide mechanisms that

explain how this state happens. In its pure form, Gestalt theory rejects the problem. It rejects the assumption that perceptions are produced out of smaller units: experience simply comes in this form. However, one can think of a Gestalt as a preferred end state: "perceptually organized" as the end-state. And this would imply that the mind has self-organizing tendencies and therefore mechanisms that do the organizing.

We get, then, two interpretations. One is explanatory and teleological. The "laws" of Gestalt psychology, which specify the basic idea of the law of *Pragnanz*, including proximity, similarity, closure, symmetry, Good Gestalt, and so forth, are descriptions of the preferred end state to which the mind tends when organizing the perceptual field out of something unorganized, such as what William James called the "great buzzing, blooming confusion" of raw percepts. Identifying the preferred end state and explaining in terms of a supposed tendency to produce it is not enough and is merely circular in the way functionalism without feedback mechanisms normally is. To turn it into a fully teleological account would require explaining the tendency, not merely citing it. So as a teleological account it is an incomplete theory, and one that is perhaps impossible to complete on its own terms. The alternative is to claim that the theory is purely descriptive, inasmuch as it does not explain why this end-state is preferred. What is omitted, in this case, is not a part of the explanation, but explanation itself.

As an incomplete teleological explanation, Gestalt theory is like local functionalism, in its traditional and Left versions, and many other social science models as well. The benefit of claiming only to be "local" and concrete is that the account is freed of the obligation to either connect to a hierarchy of ends or to be explicable in terms of non-local causes or to explain why it applies in this situation. "Incomplete" or "local," however, imply "merely descriptive" rather than explanatory. And here the issue with Gestalt theory is revealing. Gestalt theorists, like Martin, fall into the ambiguity we first encounter in Merton: they want to be explanatory, but avoid the normal means (or elements necessary to the normal means) of explanation, and provide results which are descriptive–in the case of Merton the claim of efficiency, for example–but which are supposed to be taken as explanations. The Gestaltists would reject the "teleological" account of terms like "organization" I

have supplied on their behalf, and would reject the idea that perceptual experience is built up from smaller perceptual units. They wish to have it both ways. Martin provides a sociological variant of this kind of argument. It involves two steps. The first is to reject any distinction between the phenomenologically experienced world and the objective world (presumably of causal relations):

> … the principles of field theory, and hence almost all field theorists, reject the assumption that there is in principle a divorce of phenomenology from objective description … Indeed, it was a conviction that the phenomenologically perceived world should be treated as valid that led to the emergence of Gestalt psychology and the first field investigations. In Köhler's words, the world we inhabit is 'so absolutely objective … that for a more objective world no place [is] left.' (Martin 2003: 38)

For the Gestaltists, this was why there was no point in talking about organizing experience out of smaller perceptual units: the perceived world was the one they were describing, and there was nothing "objective" behind it which the mind organized.

For Martin, "percepts are fundamentally organized because they come from a world or environment with its own principles of organization" (2003: 39). The mind does not organize: the world is pre-organized for it. This inverts the standard story about the mind full of presuppositions organizing the world that one finds in Neo-Kantianism and its heirs, such as Durkheim and Clifford Geertz. These accounts involve agency: it is the minds that are doing the organizing. Martin argues agency itself can be accounted for by field theory:

> … we directly retrieve from the environment an imperative for action: get away. There is no need to categorize and reflect; the visceral impulse to move away is anterior to the cognitive understanding. Consciousness, to take a term from Whitehead and Mead, is a 'prehensive' faculty in that the concepts it creates are grasped from the nature of the surrounding world; we are best off considering these action imperatives qualities of the

> (phenomenological) objects themselves. Similarly, our understanding of the social field is both direct and couched in the socially constructed categories that are relevant for our conduct. Consequently, field theory implies an understanding of social perception that flies in the face of the dominant pseudo-Durkheimian orthodoxy (exemplified by Mary Douglas) that sense data come in relatively disorganized form and are sorted according to a culture that, in James's words, throws 'categories over them like a net.' (Martin 2003: 39)

It is puzzling that the categories can be both "socially constructed" but not "constructed" out of something less organized: that would seem to be the very meaning of "construction." But this is the same ambiguity between "organized" and "given" that Gestaltists trade on.[4]

A field, on Martin's account, is a set of locations, pre-organized into a perceptual world consisting of affordances. "Affordance" is a term from Gibsonian psychology meaning this: some things come to us with, so to speak, handles on them, so that we would normally use them in a certain way, namely grabbing by the handles. This translates nicely into descriptions of the social world, and to descriptions of the differences between locations. To be located in a certain place in the social "field" is to have access to a distinctive set of handled things: a social role, for example the role of a physician, makes some things accessible and easy for you, and others not so easy. The social location of a ghetto drug dealer provides an entirely different set of affordances: crime is easy; going to medical school is very hard. Creating a role from scratch, or interacting without a role, is difficult.

This is a genuinely valuable use of description in aid of explanation. But the explanatory part is still murky. Martin concedes as much when he says that

> … field theory makes the exciting, nontrivial, and generative claim that action can be explained by close attention to field position as every position in the field induces a set of motivations that are subjectively experienced as 'what should be done.' Field theory disappoints us in remaining vague as to precisely

> how this occurs and we hope that it can be eventually surpassed in this regard. Yet it promises the chance of combining rigorous analytic insight with attention to the concrete. (2003: 42)

"Affordance" is an environment term: handles are inert. So something more is needed to "explain" action: "Induces" is just a vague casual term. How a motivation is induced from an external source is a mystery. Martin suggests that there is a solution to this within field theory:

> … field theorists assume that the field is defined by certain common primary motivations—subjective representations of 'what is good to strive for' such as Bourdieu's 'libido' or Emery and Trist's 'values'—and organizes other ones. (2003: 37)

But these primary motivations don't seem like the sort of things that are induced. Elsewhere he suggests that regularities such as the institutional regularity of marriage can be understood as involving "an internalization of field position" (2003: 41). Values can be understood as being subjectively experienced as injunctions, but this is "a cognitive simplification of what is otherwise a complex task of navigating a field" (2003: 37).

These kinds of claims serve to break down any distinction between the internal mental life of the individual as, for example, rational agent, and "action … understood as a field effect":

> Not only is it mistaken to assume that people choose their allies and tactics, in many cases there is no reason to assume that they choose their goals. This is not because, as critics of Bourdieu have sometimes assumed, field theory implies that people are cognitively limited in their vision of what they want. It is because the only way to reach conditions that we cognize and wish for is to make use of those conditions that we have not wished for. (2003: 44)

Cognition itself gets this treatment. Field theory gives us an

> … understanding of cognition as a seamless web of extero- and interospecific information about an intersubjectively valid field allows field theory not only to explain the correspondence of social thought to social location (previous answers in terms of interests being psychologically implausible) but to open the way for a more general understanding of the relation between subjective understandings and trajectory through a social environment. (2003: 40)

The strategy, in short, is to attribute as much as possible to facts about the field, i.e., descriptive facts, and to treat them as explanations.

This is a doomed project, if the goal is to transform field descriptions into complete explanations. Affordances and field positions are still inert: as Martin says, they are "conditions." Assumptions about generalized motivations help to complete the explanation, but they are not enough. The problem is explaining how things move from the inert facts of the field to the acting individual. Pseudo-explanatory language, such as "internalize field position" does not help. Terms like this are only placeholders for actual processes. The processes of internalization are real cognitive processes in active brains: habitualization, perceptual processing, conscious and tacit cognizing, and so forth. Of course, much can be explained by the learning that comes from navigating a field from a particular position. But "learning" and other processes beyond generalized motivations are necessary to complete the explanation. Affordances, in any case, cannot explain on their own. A handle is a handle only to the agent with a hand. The properties of the agent are part of the two way relationship that makes an affordance an affordance in the first place.

Much of Martin's discussion involves the analogy of agonistic field relations to games. Game analogies are classic "apply when they apply" cases. But Martin seems to think that the limitations of these analogies can be overcome by adding the notion that one may talk about the rules of the game and change or violate them. He commends Goffman with the comment that

> only Goffman had an orientation to norms that came anywhere near to capturing the obvious truths that conformity or

> nonconformity are strategic options with different advantages. (2003: 32)

This gives the game away: "strategic options" are the kinds of things that rational decision-makers and creative thinkers are concerned with. The addition of the idea that one is talking about the game and changing it through doing so makes it clear that the "game" is. The game is a product and causal consequence of individual action and belief, and continues to depend on these causes.

THE LIMITATIONS OF INSTRUMENTAL RATIONALITY

We can take an important lesson from field theory: location makes a difference in what people experience, what they will find convenient to believe and do, and what will work for them, and this will shape their goals and beliefs, for example in ways familiar from such works as Elster's *Sour Grapes* (1985). But it is important to keep in mind that "field theories" that describe particular fields are models of an underlying reality. The real field as opposed to the model consists of affordances, including the affordances given by the properties of other people, such as the fact that they have role ideas. The model is a description and simplification of this field.

A few comments about this notion of field are in order. Rational choice theory, formal social epistemology, and analytic sociology work with simplifying assumptions of a different kind than the assumptions made when a model of a field is constructed. They may assume omniscience or perfect knowledge. When they do they stand at the opposite poll from the description here. This account takes its cue from feminist epistemology, and the idea that knowledge is always situated, and that all situations limit and enable knowing in different ways, with different results in what people believe or how they weight these beliefs. A field, from an epistemic point of view, is a field of epistemic opportunity and limitation. We are limited to what we can learn by navigating it. These limitations form a large part of the limitations of instrumental rationality: we are not only not omniscient about the future, but

limited by our experiences and the way these experiences shape future experiences.

How does this picture fit with the explanatory claims of functionalism and field theories? If we add to our explanatory tool kit some considerations about the limitations of instrumental rationality, of the sort discussed by Cherkaoui (2007: 56-57), we can provide an alternative. Obviously there are many other considerations that can also be applied, but in what follows I will concentrate on this one. My concern will be to show that it is possible to answer explanatory questions using the notion of field as I have outlined it, without appealing to occult forces or processes, like "internalization of field position" or mysterious collective objects like *habitus,* and to show that the same kinds of considerations about the limits of instrumental reason allow us to account for the kinds of things functionalism attempts to account for.

A classic example is given by Malinowski (1944). In the Trobriand Islands there were two kinds of fishing. Lagoon fishing was predictable and safe, and done by poisoning; open sea fishing depended on the presence of schools of fish, and was risky, full of danger and uncertainty. In the case of sea fishing, but not lagoon fishing, there was "extensive magical ritual." This was ritual with a specific aim, which the natives could all articulate. The beliefs in the effects of these rituals were false, of course. The real effect of these rituals, according to Malinowski, was that they reduced anxiety, anxiety well justified by the risks of sea fishing, risks absent in the case of lagoon fishing, where there were no rituals. Misrecognition is central to these rituals: the Trobrianders think they are to protect them from specific dangers. The effect is to reduce anxiety. If they knew the rituals had no effect on the dangers, they would no longer function to reduce anxiety.

Elster's question immediately comes to mind: in what sense can the function of reducing anxiety be said to cause the persistence of the use of the ritual without recognition of this effect? One can imagine a variety of answers to this. One would be that they do recognize the effect, and this is one of their motives for performing the ritual in addition to the main motive of preserving against danger. Another would be that they unconsciously appreciate the effect even though they could not articulate the causal connection between the ritual and the effect, and

are thus conditioned to perform the ritual. This would be Malinowski's view: anxiety was for him a hidden cause that played a role in magical ritual generally, and not a native's concept. Both of these are feedback accounts. Evidence is non-existent in either case: Malinowski acknowledges that he is dealing in abstractions, not observational data.

A simpler explanation can be given in terms of the issues of uncertainty that Malinowski himself stresses, in terms of the limitations of instrumental rationality. The rituals are clearly instrumental in their intention. They are also rational. The rituals reflect a Pascalian wager that is very cheap. It costs little to act on the belief that the ritual does some good; it potentially costs a great deal to wrongly disbelieve in the efficacy of the ritual. On a purely individual basis, it would be good, under these conditions of uncertainty, to perform the ritual. When others are involved, it becomes better: one avoids the inconvenience of explaining why one is violating the ritual, potentially endangering others, and of explaining why one is rejecting what everyone else appears to believe. Radcliffe-Brown made similar points about this explanation in the course of supplying an explanation of similar rituals preserving a woman from the dangers of childbirth:

> In a given community it is appropriate that an expectant father should feel concern or at least make an appearance of doing so. Some suitable symbolic expression of his concern is found in terms of the general ritual or symbolic idiom of the society, and it is generally felt that a man in that situation ought to carry out the symbolic or ritual actions or abstentions. (1939: 41)

This appeals to field position, that of the expectant father, and the affordances available to him, namely the ritual idiom of the society, and the expectations of others, which provide a motive, taken together with his own desires for a convenient solution, a convenient thing to believe or do that would avoid the consequences attendant on failing to meet expectations.

Nothing functionalist figures in these explanations: there is no occult force, or hypothesized feedback loop involving an abstract emotional state. In context, these are powerful reasons to believe. And the

context is one of cognitive limitations and specific affordances: the readily available beliefs of others and the consistency of these beliefs with available ritual practices. Is there anything left over? A fully elaborated account might include the tacit processes that produce habits and the cognitive preconditions for understanding. But none of these involve occult processes. There is nothing left over that requires unusual theoretical objects and the unconscious feedback processes that might fit with the kinds of functionalist explanations that normally figure in social institutions will be minor causes compared to those discussed here. Obviously there is a level of explanation that could involve deep psychic forces, in which there is feedback in which something like anxiety is relieved and unconsciously reinforced. But it is implausible to think that this feedback loop would overwhelm all these other considerations and be the dominant cause.

A FINAL METHODOLOGICAL COMMENT

Social science models are abstractions and simplifications. But it is well not to forget the underlying reality which models are simplifying. If one considers the reality I have outlined above, in which knowledge and tacit or habitualized knowledge varies in complex ways between individuals with experiences conditioned by different social locations and environments, this much is clear: there is no prospect of a "rigorous," meaning mathematically tractable, model that corresponds more than very roughly to this reality, or even to aspects of this reality. But simplifications for a purpose, such as prediction, can work and be revealing. Balance of power theory in international relations, for example, has many exceptions, but it is robust and predictive over a wide range of cases of great power conflict (Mearsheimer 2001: 29-54). It is nevertheless limited. As Hans Morgenthau puts it, "political reality is replete with contingencies and systematic irrationalities" (1978: 8). And few models are as successful as this one.

Claims to have captured the essence of some complex social phenomenon need to be taken *cum grano salis*. Producing the appearance of doing so normally involves sleight of hand. Malinowski focused his readers on a specific contrast: between lagoon and sea fishing, in terms

of which one specific difference, danger, stood out. The anxiety theory of ritual which was built on this case was, however, a grotesque simplification. It elevated a small feedback loop to a major cause, and then to an essence. This is unfortunately not an isolated case. There is a symbiotic relation between claims about essences and exotic theoretical objects. If one thinks one has the essence of something: ritual, or schooling, it is all the more tempting to invent a hidden force or secret purpose that explains this essence. And if one thinks there are secret forces of oppression or system maintenance it is all the more tempting to think one has captured an essence. But the long history of these claims makes this much clear: the illumination of social science models is only partial, and often very temporary. And the exotic theoretical objects that are invoked in these models rarely outlive the moment for which the model was constructed.

Use English translation as reference: *Good Intentions: The Paradox of Unintended Concesquences.* Oxford: The Bardwell Press, 2007?

NOTES

1. For the generic problems with the kinds of explanatory devices used by Bourdieu and others, which I call collective objects, see Turner (1994).
2. Martin says a great deal about how field theories are "intrinsically at odds" with "mechanistic explanation" (2003: 10), and claims that "social fields differ from board games, however, in that the struggle is both over and within the rules, and thus the 'game' is not some sort of overarching formal framework that in some obscure way forces persons to do this or that." Nevertheless, he uses familiar causal language: "fields are only known by their effects." When he describes the *habitus* as "a cultural unconscious, a matrix of dispositions that serves to affectively organize perceptions" (Bourdieu 1969: 182 cited in Martin 2003: 23), he uses classically "causal" language: "dispositions" are causes, and they do something to people: to "organize" their perceptions must "in some obscure way" involve "force." Why these causal usages would not require "mechanisms" to be substantiated, or why Martin thinks he can divorce the language of effect from that of cause, is obscure.
3. Martin says the following: "… field theories, like mechanistic theories (and unlike functionalist theories), reach toward the concrete and propose only local action, but like functionalist theories (and unlike mechanistic theories), they insist that any case must be understood in terms of the global pattern" (2003: 12). What can "local action" mean other than "action" involving some sort of mechanism, causal or teleological? And what is a global pattern other than the model (which can potentially be applied elsewhere, but may not apply) itself? As

we will see, Martin thinks he can answer this question in terms of the notion of affordances, which he takes to be descriptive givens of the situation which "produce" various outcomes non-causally or at least in terms other than those of causal mechanisms of the usual kind.

4. This is precisely the issue seized on by David Hamlyn in his classic critique of the explanatory claims of Gestalists (1958). Hamlyn's point is that there can be no explanation without a contrast class. In this case the contrast would be to the pre-organized world. But if we deny there is a pre-organized world, there is no difference to explain.

REFERENCES

Bourdieu, Pierre ([1997] 2000) *Pascalian Meditations*. Oxford, Polity Press.

Cherkaoui, Mohamed (2007) *Good Intentions: Max Weber and the Paradox of Unintended Consequences*, Trans. Peter Hamilton. Oxford: Bardwell Press.

Douglas, Mary (1986) *How Institutions Think*. Syracuse, NY, Syracuse University Press.

Elster, Jon (1985) *Sour Grapes: Studies in the Subversion of Rationality.* Cambridge: Cambridge University Press.

——— (1990) "Merton's functionalism and the unintended consequences of action", in Jon Clark, Celia Modgil and Sohan Modgil (eds), *Robert Merton: Consensus and Controversy*. London/New York: Falmer Press: 129–135.

——— (2010) "Obscurantisme dur et obscurantisme mou dans les sciences humaines et sociales", *Diogène,* 1: 231–147.

Emery, F.E., and Trist, E.L. (1965) "The causal texture of organization environments", *Human Relations,* 18: 21–32.

Goffman, Erving. 1959. *The Presentation of Self in Everyday Life*. New York, Anchor Books.

Hamlyn, David W. (1951) "Psychological explanation and the Gestalt hypothesis." *Mind*, 60(240): 506-520.

James, William ([1909] 1975) *The Meaning of Truth*. Cambridge, MA: Harvard University Press.

Koffka, K. (1935) *Principles of Gestalt Psychology*. New York: Harcourt, Brace.

Malinowski, Bronislaw (1944) *A Scientific Theory of Culture and Other Essays*. Chapel Hill, NC: The University of North Carolina Press.

Martin, John Levi (2003) "What is field theory?" *American Journal of Sociology,* 109(1): 1–49.

Mead, George H. (1926) "The objective reality of perspectives." *Proceedings of the Sixth International Congress of Philosophy,* 6: 75–87.

Mearsheimer, John J. (2001) *The Tragedy of Great Power Politics*. New York: Norton.

Merton, Robert (1958) "The functions of the professional association." *The American Journal of Nursing,* 58(1): 50-54.

Morgenthau, H. ([1948] 1978) *Politics Among Nations: The Struggle for Peace and Power*, 5th rev edn. New York: A. A. Knopf.

Radcliffe-Brown, A.R. (1939) *Taboo: The Frazer Lecture*. Cambridge: Cambridge University Press.

Turner, Stephen (1994) *The Social Theory of Practices: Tradition, Tacit Knowledge, and Presuppositions*, Oxford: Polity Press; Chicago: University of Chicago Press.
Weber, Max ([1930] 1958) *The Protestant Ethic and the Spirit of Capitalism*, trans. Talcott Parsons. New York: Charles Scribner's Sons.
Whitehead, Alfred North ([1929] 1978) *Process and Reality: An Essay in Cosmology*. New York: Free Press.