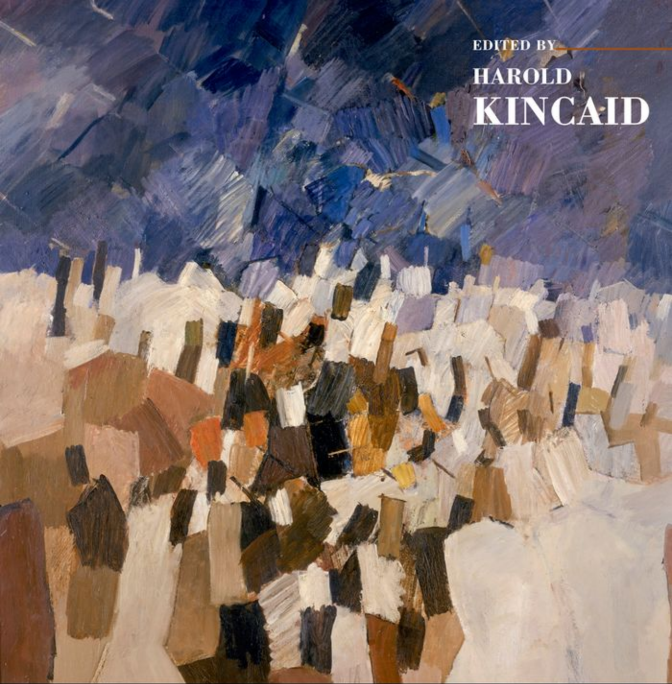
EDITED BY
HAROLD KINCAID



The Oxford Handbook *of* PHILOSOPHY OF SOCIAL SCIENCE

THE OXFORD HANDBOOK OF

PHILOSOPHY OF SOCIAL SCIENCE

This page intentionally left blank

THE OXFORD HANDBOOK OF

PHILOSOPHY OF SOCIAL SCIENCE

Edited by

HAROLD KINCAID

OXFORD
UNIVERSITY PRESS

OXFORD
UNIVERSITY PRESS

Oxford University Press is a department of the University of Oxford.
It furthers the University's objective of excellence in research, scholarship,
and education by publishing worldwide.

Oxford New York
Auckland Cape Town Dar es Salaam Hong Kong Karachi
Kuala Lumpur Madrid Melbourne Mexico City Nairobi
New Delhi Shanghai Taipei Toronto

With offices in
Argentina Austria Brazil Chile Czech Republic France Greece
Guatemala Hungary Italy Japan Poland Portugal Singapore
South Korea Switzerland Thailand Turkey Ukraine Vietnam

Published in the United States of America by
Oxford University Press
198 Madison Avenue, New York, NY 10016

www.oup.com

Library of Congress Cataloging-in-Publication Data
The Oxford handbook of philosophy of social science /edited by Harold Kincaid.
p. cm.—(Oxford handbooks)
Includes bibliographical references.
ISBN 978-0-19-539275-3 (alk. paper)
1. Social sciences—Philosophy. 2. Social sciences—Research.
I. Kincaid, Harold, 1952-II. Title: Philosophy of social science.
H61.O95 2012
300.1—dc23 2011036789

1 3 5 7 9 8 6 4 2

Printed in the United States of America
on acid-free paper

Dedicated to the intercontinental sundowners support group:
Nelleke Bak
Robin Lessel
Don Ross

This page intentionally left blank

Table of Contents

Preface xi
Contributors xiii

1. Introduction: Doing Philosophy of Social Science 3
Harold Kincaid

Part I: Mechanisms, Explanation, and Causation

2. Micro, Macro, and Mechanisms 21
Petri Ylikoski

3. Mechanisms, Causal Modeling, and the Limitations of Traditional Multiple Regression 46
Harold Kincaid

4. Process Tracing and Causal Mechanisms 65
David Waldner

5. Descriptive-Causal Generalizations: "Empirical Laws" in the Social Sciences? 85
Gary Goertz

6. Useful Complex Causality 109
David Byrne and Emma Uprichard

7. Partial Explanations in Social Science 130
Robert Northcott

8. Counterfactuals 154
Julian Reiss

9. Mechanistic Social Probability: How Individual Choices and Varying Circumstances Produce Stable Social Patterns 184
Marshall Abrams

Part II: Evidence

10. The Impact of Duhemian Principles on Social Science Testing and Progress 229
Fred Chernoff

11. Philosophy and the Practice of Bayesian Statistics in the Social Sciences 259
Andrew Gelman and Cosma Rohilla Shalizi

12. Sciences of Historical Tokens and Theoretical Types: History and the Social Sciences 274
Aviezer Tucker

13. RCTs, Evidence, and Predicting Policy Effectiveness 298
Nancy Cartwright

14. Bringing Context and Variability Back into Causal Analysis 319
Stephen L. Morgan and Christopher Winship

15. The Potential Value of Computational Models in Social Science Research 355
Ken Kollman

Part III: Norms, Culture, and the Social-Psychological

16. Models of Culture 387
Mark Risjord

17. Norms 409
David Henderson

18. The Evolutionary Program in Social Philosophy 436
Francesco Guala

19. Cultural Evolution: Integration and Skepticism 458
Tim Lewens

20. Coordination and the Foundations of Social Intelligence 481
Don Ross

21. Making Race Out of Nothing: Psychologically Constrained Social Roles 507
Ron Mallon and Daniel Kelly

Part IV: Sociology of Knowledge

22. A Feminist Empirical and Integrative Approach in Political Science: Breaking Down the Glass Wall? 533
Amy G. Mazur

23. Social Constructions of Mental Illness 559
Allan Horwitz

Part V: Normative Connections

24. Cooperation and Reciprocity: Empirical Evidence and Normative Implications 581
James Woodward

25. Evaluating Social Policy 607
Daniel M. Hausman

26. Values and the Science of Well-Being: A Recipe for Mixing 625
Anna Alexandrova

Index 647

This page intentionally left blank

Preface

This volume results from a collaborative effort in several respects. All but two of the chapters were presented in draft form at a conference at the University of Alabama at Birmingham in 2010. The contributors made for a lively and thoughtful audience, and I am sure their comments at the conference substantially improved the chapters. After written drafts were submitted, each contributor commented on one or two other contributions in some detail. In addition, Steve Morgan, Christopher Winship, Don Ross, Gary Goertz, and Aviezer Tucker all provided useful comments on the initial proposal that led to significant improvements.

This topics represented in this handbook are shaped by several things. This volume is preceded by another: H. Kincaid and D. Ross, eds., *Oxford Handbook of the Philosophy of Economics* (Oxford: Oxford University Press, 2009). As a result, little discussion of economics has been included in this volume. The aim of the volume was to promote philosophy of science in the naturalist vein that engaged with ongoing current controversies in social research (as explained and defended in the introduction), and the chapters included strongly reflect that goal.

I want to thank UAB's Center for Ethics and Values in the Sciences for support in organizing the conference for the volume, and Peter Ohlin at Oxford University Press for encouragement and advice.

This page intentionally left blank

Contributors

Marshall Abrams is assistant professor in the Department of Philosophy at the University of Alabama at Birmingham. He received his PhD from the University of Chicago and was an NSF-sponsored postdoctoral fellow at Duke University's Center for Philosophy of Biology. His philosophical research focuses on the nature and role of probability and causation in evolutionary biology and the social sciences, and on interactions between biological evolution and social processes with emphasis on modeling of cognitive coherence relations in cultural change. He is also engaged in purely scientific research on the evolution of obesity and diabetes, and is an associate editor at the journal *Frontiers in Evolutionary and Population Genetics*.

Anna Alexandrova is a philosopher of social science at Cambridge University. She has taught at University of Missouri, St. Louis, and received her PhD from the University of California, San Diego. She has written on the use of formal models for explanation and policy making in economics and history, and on the measurement of happiness in psychology. Her current work examines well-being as an object of science and a social indicator.

David Byrne is professor of sociology and social policy in the School of Applied Social Sciences, Durham University. He has worked at the interface between the academy and the application of social science throughout his career, including a period as research director of a community development project. Publications include *Beyond the Inner City* (Open University Press, 1989), *Complexity Theory and the Social Sciences* (Routledge, 1998), *Social Exclusion* (Open University Press, 2009), *Understanding the Urban* (Palgrave-Macmillan, 2001), *Interpreting Quantitative Data* (Sage, 2002), and *Applying Social Science* (Policy Press, 2011).

Nancy Cartwright is professor of philosophy at the Department of Philosophy, Logic and Scientific Method at the London School of Economics and Political Science and at the University of California, San Diego. She was president of the Philosophy of Science Association in 2010 and president of the American Philosophical Association (Pacific Division) in 2008. Her research interests include philosophy and history of science (especially physics and economics), causal inference, and objectivity and evidence, especially on evidence-based policy. Her publications include *How the Laws of Physics Lie* (Oxford University Press, 1983), *Nature's Capacities and Their Measurement* (Oxford University Press, 1989), *Otto Neurath: Philosophy between Science and Politics* (Cambridge University Press, 1995, with Jordi Cat, Lola Fleck, and Thomas E. Uebe), *The Dappled World: A Study of the*

Boundaries of Science (Cambridge University Press, 1999), and *Hunting Causes and Using Them* (Cambridge University Press, 2007).

Fred Chernoff is Harvey Picker Professor of International Relations and chair of the Department of Political Science at Colgate University. He holds a PhD in philosophy from Johns Hopkins University and a PhD in political science from Yale University. He has held research posts at the International Institute for Strategic Studies, the Rand Corporation, and the Norwegian Institute of International Affairs. He is author of *After Bipolarity* (University of Michigan Press, 1995), *The Power of International Theory* (Routledge, 2005) and *Theory and Metatheory in International Relations* (Palgrave-Macmillan, 2007) and has contributed to many journals of international relations and philosophy, including *International Studies Quarterly*, *Journal of Conflict Resolution*, *Millennium*, *International Theory*, *European Journal of International Relations, Philosophical Quarterly, Mind*, and *Analysis*.

Andrew Gelman is a professor of statistics and political science and director of the Applied Statistics Center at Columbia University. His books include *Bayesian Data Analysis* (Chapman and Hall, 2003, with John Carlin, Hal Stern, and Don Rubin), *Teaching Statistics: A Bag of Tricks* (Oxford University Press, 2002, with Deb Nolan,), *Data Analysis Using Regression and Multilevel/Hierarchical Models* (Cambridge University Press, 2006, with Jennifer Hill), and, most recently, *Red State, Blue State, Rich State, Poor State: Why Americans Vote the Way They Do* (Princeton University Press, 2009, with David Park, Boris Shor, Joe Bafumi, and Jeronimo Cortina). Andrew has done research on a wide range of topics, including why it is rational to vote, why campaign polls are so variable when elections are so predictable, why redistricting is good for democracy, reversals of death sentences, police stops in New York City, the statistical challenges of estimating small effects, the probability that your vote will be decisive, seats and votes in Congress, and social network structure.

Gary Goertz is professor of political science at the University of Arizona. He is the author or editor of nine books and over forty articles on issues of methodology, international institutions, and conflict studies, including *Necessary Conditions: Theory, Methodology, and Applications* (Rowman & Littlefield, 2003), *Social Science Concepts: A User's Guide* (Princeton University Press, 2006), *Explaining War and Peace: Case Studies and Necessary Condition Counterfactuals* (Routledge, 2007), *Politics, Gender, and Concepts: Theory and Methodology* (Cambridge University Press, 2008), and *A Tale of Two Cultures: Contrasting Qualitative and Quantitative Paradigms* (Princeton University Press, 2012).

Francesco Guala is associate professor of economics and philosophy at the University of Milan, Italy. His research focuses on methodological and ontological problems arising from the social sciences. He is the author of *The Methodology of Experimental Economics* (Cambridge University Press, 2005) and coeditor, with Daniel Steel, of *The Philosophy of Social Science Reader* (Routledge, 2011).

Daniel M. Hausman is the Herbert A. Simon and Hilldale Professor of Philosophy at the University of Wisconsin-Madison. His research has centered on epistemological, metaphysical, and ethical issues lying at the boundaries between economics and philosophy. His books include *The Inexact and Separate Science of Economics* (Cambridge University Press, 1992) and *Economic Analysis, Moral Philosophy, and Public Policy* (Cambridge University Press, 2006, with Michael McPherson). His most recent book, *Preferences, Value, and Choice*, *and Welfare*, is due out shortly from Cambridge University Press.

David Henderson is the Robert R. Chambers Professor of Philosophy at the University of Nebraska, Lincoln. He has written on interpretation and explanation in the social sciences, with special concern for the place for finding rationality in those matters. He also writes in epistemology, where he is interested in the epistemological implications of recent work in cognitive science.

Allan Horwitz is Board of Governors Professor of Sociology at Rutgers University. His most recent books are *The Loss of Sadness: How Psychiatry Transformed Normal Sorrow into Depressive Disorder* (Oxford University Press, 2007, with Jerome Wakefield), *Diagnosis, Therapy, and Evidence: Conundrums in Modern American Medicine* (Rutgers University Press, 2010, with Gerald Grob), and *All We Have to Fear: Anxiety and the Boundaries of Normality* (Oxford University Press, 2012, with Jerome Wakefield).

Daniel Kelly is an assistant professor in the philosophy department at Purdue University. His research interests are at the intersection of the philosophy of mind, cognitive science, and moral theory. He is the author of *Yuck! The Nature and Moral Significance of Disgust* (The MIT Press, 2011), and has published papers on moral judgment, social norms, racial cognition, and cross-cultural diversity.

Harold Kincaid is professor in the School of Economics at the University of Cape Town, South Africa. He is the author of *Philosophical Foundations of the Social Sciences* (Cambridge University Press, 1996) and *Individualism and the Unity of Science* (Rowman & Littlefield, 2007), and coeditor of *What Is Addiction?* (The MIT Press, 2009), *The Oxford Handbook of Philosophy of Economics* (Oxford University Press, 2009), and three other volumes. He has published widely in the philosophy of the social sciences and philosophy of economics. His current research interests include causal modeling and empirical studies of addiction and time and risk attitudes in the developing world.

Ken Kollman is Frederick G. L. Huetwell Professor and Professor of Political Science at the University of Michigan, Ann Arbor. His research and teaching focus on political parties, elections, lobbying, federal systems, formal modeling, and complexity theory. In addition to numerous articles, he has written *The Formation of National Party Systems: Federalism and Party Competition in Canada, Great Britain, India, and the United States* (Princeton University Press, 2004, with Pradeep Chhibber), *Outside Lobbying: Public Opinion and Interest Group Strategies* (Princeton University Press, 1998), and *The American Political System* (Norton, 2011).

Tim Lewens is reader in philosophy of the sciences at the University of Cambridge, where he is also a fellow of Clare College. His previous publications include *Darwin*

(Routledge, 2007) and *Organisms and Artifacts: Design in Nature and Elsewhere* (The MIT Press, 2004).

Ron Mallon is an associate professor of philosophy and director of the Philosophy-Neuroscience-Psychology Program at Washington University in St. Louis. His research is in social philosophy, philosophy of cognitive psychology, and moral psychology. He has authored or coauthored papers in *Cognition*, *Ethics*, *Journal of Political Philosophy*, *Midwest Studies in Philosophy*, *Mind and Language*, *Noûs*, *Philosophy and Phenomenological Research*, *Philosophy of Science*, *Social Neuroscience, Social Philosophy*, and *Social Theory and Practice*.

Amy G. Mazur is professor of political science at Washington State University. She is coeditor of *Political Research Quarterly*. Her recent publications include *Politics, Gender and Concepts* (edited with Gary Goertz, Cambridge University Press, 2008) and *The Politics of State Feminism: Innovation in Comparative Research* (Temple University Press, 2010, with Dorothy McBride).

Stephen L. Morgan is professor of sociology and the director of the Center for the Study of Inequality at Cornell University. He has a PhD in sociology from Harvard University and an MPhil in comparative social research from Oxford University. He has published two books: *On the Edge of Commitment: Educational Attainment and Race in the United States* (Stanford University Press, 2005) and, coauthored with Christopher Winship, *Counterfactuals and Causal Inference: Methods and Principles for Social Research* (Cambridge University Press, 2007).

Robert Northcott is currently a lecturer in philosophy at Birkbeck College in London. Before that, he was assistant professor at the University of Missouri-St. Louis. He has published widely on, among other things, causation, causal explanation, and degree of causation, including on how these notions are conceptualized and used in social science.

Julian Reiss is associate professor in the philosophy faculty of Erasmus University Rotterdam, and specializes in philosophy of economics and general philosophy of science. Specific research interests are causal inference, measurement, models and thought experiments, and the place of values in science. Publications include *Error in Economics: Towards a More Evidence-Based Methodology* (Routledge, 2008), *Causality Between Metaphysics and Methodology* and *Philosophy of Economics* (both forthcoming with Routledge), and thirty-five papers in journals such as *Philosophy of Science*, *Synthese*, *Philosophy of the Social Sciences*, and *Theoria*.

Mark Risjord is a professor of philosophy at Emory University. He received his PhD in philosophy from the University of North Carolina, Chapel Hill in 1990. His research interests include the epistemological foundations of the social sciences, the role of values in scientific research, and the philosophy of the health sciences.

Don Ross is professor of economics and dean of commerce at the University of Cape Town, and a research fellow in the Center for Economic Analysis of Risk at Georgia State University. His main areas of research include the experimental economics of nonstandard consumption patterns, the philosophical foundations of economics and game theory, naturalistic philosophy of science, and trade and industry policy in Africa. He is the author of numerous articles and books, including *Economic Theory and Cognitive Science: Microexplanation* (The MIT Press, 2005) and *Every Thing Must Go: Metaphysics Naturalized* (Oxford University Press, 2007, with James Ladyman). He is coeditor (with Harold Kincaid) of *The Oxford Handbook of Philosophy of Economics* (Oxford University Press, 2009).

Cosma Rohilla Shalizi is an assistant professor of statistics at Carnegie Mellon University and an external professor at the Santa Fe Institute. He received his PhD in theoretical physics from the University of Wisconsin-Madison in 2001. His research focuses on time series prediction, network analysis, and inference in complex systems.

Aviezer Tucker is the author of *Our Knowledge of the Past: A Philosophy of Historiography* (Cambridge University Press, 2004) and the editor of *The Blackwell Companion to the Philosophy of History and Historiography* (Wiley-Blackwell, 2009). His research concentrates on the philosophy of the historical sciences, epistemology, and political philosophy. He lives in Austin, Texas. Previously he lived in New York, Prague, Canberra, and a few other places.

Emma Uprichard is a senior lecturer in the Department of Sociology, Goldsmiths, University of London. She is particularly interested in the methodological challenge of applying complexity theory to the study of change and continuity in the social world. She has substantive research interests in methods and methodology, critical realism, cities, time and temporality, children and childhood, and food.

David Waldner is associate professor of politics at the University of Virginia, where he teaches courses on the political economy of developing nations and methodology. He is the author of *State Building and Late Development* (Cornell University Press, 1999) and is currently writing two books, *Democracy & Dictatorship in the Post-Colonial World* and *Causation, Explanation, and the Study of Politics*.

Christopher Winship is the Diker-Tishman Professor of Sociology at Harvard University and a member of the Harvard Kennedy School of Government's senior faculty. Prior to coming to Harvard in 1992, he was a professor in sociology, statistics, and (by courtesy) economics. At Harvard he is a member of the criminal justice program, inequality program, and the Hauser Center for the Study of Nonprofits. His research interests include models of selection bias, causality, youth violence, pragmatism, and the implications of the cognitive revolution for sociology. With Stephen Morgan, he is author of *Counterfactuals and Causal Inference* (Cambridge, 2007).

James Woodward is distinguished professor in the Department of History and Philosophy of Science at the University of Pittsburgh. He was formerly the J. O. and Juliette Koepfli Professor of the Humanities at the California Institute of Technology. His book *Making Things Happen: A Theory of Causal Explanation* (Oxford

University Press, 2003) won the Lakatos award in 2005. He is president of the Philosophy of Science Association in 2010–12.

Petri Ylikoski is currently an academy research fellow at the University of Helsinki, Finland. His research interests include theory of explanation, science studies, philosophy of the social sciences, and philosophy of biology. He is especially interested in the interfaces between the social and the biological sciences and in the idea of mechanism-based understanding in the social sciences.

THE OXFORD HANDBOOK OF

PHILOSOPHY OF SOCIAL SCIENCE

This page intentionally left blank

CHAPTER 1

INTRODUCTION: DOING PHILOSOPHY OF SOCIAL SCIENCE

HAROLD KINCAID

This volume is shaped by important developments in both the social sciences and the philosophy of the social sciences over the last several decades. In this chapter I outline these changes and argue that they have indeed been significant advances in our thinking about the social world. Rather than providing linear summaries of twenty-plus chapters, I delineate the frameworks and issues that motivate the kind of philosophy of social science and social science that is represented in this volume. Both philosophy of social science and social science itself are intermixed in the following chapters. That is because the volume is built around a guiding naturalism that denies that there is something special about the social world that makes it unamenable to scientific investigation, and also denies that there is something special about philosophy that makes it independent or prior to the sciences in general and the social sciences in particular. In the process of outlining recent developments the chapters of the handbook are related and motivated, and open unresolved issues are discussed.

1.1. Developments in Philosophy of Science

I start with developments in the philosophy of science. Though the monikers are not entirely historically accurate, I want to contrast previous positivist philosophy of science with postpositivist views which I believe provide a much more useful framework for thinking about science and social science. Some of the key tenets of positivist philosophy of science are as follows.[1]

Theories are the central content of science. A mature science ideally produces one clearly identifiable theory that explains all the phenomena in its domain. In practice, a science may produce different theories for different subdomains, but the overarching scientific goal is to unify those theories by subsuming them under one encompassing account. Theories are composed of universal laws relating and ascribing properties to natural kinds and are best understood when they are described as formalized systems. Philosophy of science can aid in producing such formalizations by the application of formal logic.

The fundamental concepts of science should have clear definitions in terms of necessary and sufficient conditions. General philosophy of science is in large part about clarifying general scientific concepts, especially explanation and confirmation. The goal is to produce a set of necessary and sufficient conditions for application of these concepts. These definitions are largely tested against linguistic intuitions about what we would and would not count as cases of explanation and confirmation.

Explanation and confirmation have a logic—they conform to universal general principles that apply to all domains and do not rest on contingent empirical knowledge. A central goal of philosophy of science is to describe the logic of science. Explanation involves (in some sense still to be clarified) deductions from laws of the phenomena to be explained. Whether a science is well supported by evidence can be determined by asking whether the theory bears the right logical relationship to the data cited in support of it.

Independence of philosophy from science: Identifying the logic of inference and explanation and the proper definition of concepts are philosophical activities. Scientists certainly can act as philosophers, but the philosophy and the science are different enterprises with different standards. The collorary is that philosophy of science is largely done after the science is finished.

Social institutions are irrelevant. The social organization of science may be an interesting topic for sociologists, but it has little direct bearing on philosophy of science's tasks.

The criteria for explanation and confirmation allow us to properly demarcate scientific theories from pseudoscientific accounts. Pseudoscientific accounts tend to sacrifice due attention to confirmation in favor of apparent explanation, and in so doing fail to be genuinely explanatory.

It is a serious open question to what extent any of the social sciences are real sciences. This question is best explored by comparing their logical structures with those characteristic of physics and, to a lesser extent chemistry, geology, and biology. All the key characteristics described above should characterize any scientific social science and its related philosophy of science.

These positivist ideas have been replaced with a considerably more subtle and empirically motivated view of the philosophy of science in the following ways.

Theories as central: "The" theory in a given discipline is typically not a single determinate set of propositions. What we find instead are common elements that are given different interpretations according to context. For example, genes play a central role in biological explanation, but what exactly a gene is taken to be varies considerably depending on the biological phenomena being explained (Moss 2004). Often we find *no* one uniform theory in a research domain, but rather a variety of models that overlap in various ways but that are not fully intertranslatable. Cartwright (1980) gives us the example of models of quantum damping, in which physicists maintain a toolkit of six different mathematical theories. Because these aren't strictly compatible with one another, a traditional perspective in the philosophy of science would predict that physicists should be trying to eliminate all but one. However, because each theory is better than the others for governing some contexts of experimental design and interpretation, but all are reasonable in light of physicists' consensual informal conception of the basic cause of the phenomenon, they enjoy their embarrassment of riches as a practical boon. There is much more to science than theories: experimental setup and instrument calibration skills, modeling ingenuity to facilitate statistical testing, mathematical insight, experimental and data analysis paradigms and traditions, social norms and social organization, and much else—and these other elements are important to understanding the content of theories.

Theories, laws, and formalization: Laws in *some* sense play a crucial role in scientific theories. Absent any trace of what philosophers call modal structure, it is impossible to see how scientists can be said to rationally learn from induction. However, some of our best science does not emphasize laws in the philosopher's sense as elegant, context-free, universal generalizations, but instead provides accounts of temporally and spatially restricted context-sensitive causal processes as its end product. Molecular biology is a prime example in this regard, with its emphasis on the causal mechanisms behind cell functioning that form a complex patchwork of relations that cannot be aggregated into an elegant framework. Expression in a clear language—quantitative where possible—is crucial to good science, but the ideal of a full deductive system of axioms and theorems is often unattainable and not, as far as one can see, actually sought by many scientific subcommunities that are nevertheless thriving.

Conceptual analysis: Some important scientific concepts are not definable in terms of necessary and sufficient conditions but are instead much closer to the prototypes that, according to cognitive science, form the basis for our everyday concepts of kinds of entities and processes. The concept of the gene is again a good example. There is no definition of gene in terms of its essential characteristics that covers every important scientific use of the concept. Cartwright (2007) has argued recently that the same holds even for so general and philosophical an idea as *cause*: There are different senses of *cause* with different relevant formalizations and evidence conditions. Equally important, the traditional philosophical project of testing definitions against what we find it appropriate to say is of doubtful significance. Who is the relevant reference group? The intuitive judgments of philosophers, whose grasp of science is often out of date and who are frequently captured by highly specific metaphysical presuppositions, do not and should not govern scientific usage at all (Ladyman and Ross 2007, chapter 1). Questions about the usage of scientists is certainly more relevant, but this also may not be the best guide to the content of scientific results.

The logic of confirmation and explanation: Confirmation and explanation are complex practices that do not admit of a uniform, purely logical analysis. Explanations often have a contextual component set by the background knowledge of the field in question that determines the question to be answered and the kind of answer that is appropriate. Sometimes that context may invoke laws, but often it does not, at least not in any explicit way. Confirmation likewise depends strongly on domain-specific background knowledge in ways that make a purely logical and quantitatively specifiable assessment of the degree to which specified evidence supports a hypothesis unlikely. The few general things that can be said about confirmation are sufficiently abstract that they are unhelpful on their own. The statements "a hypothesis is well supported if all sources of error have been ruled out" or "a hypothesis is well supported by the evidence if it is more consistent with the evidence than any other existing hypothesis" are hard to argue with. Yet to make any use of these standards in practice requires fleshing out how error is ruled out in the specific instance or what consistency with the evidence comes to in that case. Other all-purpose criteria such as "X is confirmed if and only if X predicts novel evidence" or "X is confirmed if and only if X is the only hypothesis that has not been falsified" are subject to well-known counter examples and difficulties of interpretation.

Holism: It is a fallacy to infer from the fact that every hypothesis is tested in conjunction with background theory that evidence only bears on theories as wholes (Glymour 1980). By embedding hypotheses in differing background theoretical and experimental setups, it is possible to attribute blame and credit to individual hypotheses. Indeed, this is how the overwhelming majority of scientists view the overwhelming majority of their own research results. Judged on the basis of considerations that scientists typically introduce into actual debates about what to regard as accepted results, the relationships between theories, applications, and tests propagated by Quine, Kuhn, and Lakatos look like philosophers' fantasies. While these three philosophers were instrumental in the transition from positivist philosophy of science, their arguments and views have been superceded: Data may be theory-laden, but theory-laden comes to many things and does not mean that every piece of data is laden with whole theories, and does not prevent the kind of triangulation and piecemeal testing of specific hypotheses characteristic of good science.

Independence of philosophy from science: Philosophy of science and science are continuous in several senses. As we saw, the traditional conceptual analysis of analytic philosophy is a nonstarter and philosophical claims are subject to broad empirical standards of science. Of course, getting clear on concepts has real value. However, it is something scientists do all the time, but in ways far more sophisticated and empirically disciplined than the traditional philosophical practice of testing proposed definitions against what we would say or against intuitions (Wilson 2007). Philosophy of science is also continuous with science in that philosophy of science is not entirely or mostly something that is done after the science is settled. Instead, philosophy of science issues arise in ongoing scientific controversies and part of the process of settling those issues. Again, philosophy of science is something that scientists themselves do, and in a sense science is something that philosophers of science do. Contemporary philosophy of biology is a paradigm case in this regard. Philosophers of science publish in biology journals and biologists publish in philosophy of biology venues. The problems tackled are as much biological as philosophical or conceptual: The questions are such things as how is genetic drift to be understood or what is the evidence for group selection.

Science and pseudoscience: Several of the insights about science already discussed suggest that judging theories to be scientific or pseudoscientific is a misplaced enterprise. Scientific theories and their evidence form complexes of claims that involve diverse relations of dependence and independence and, as a result, are not subject to uniform or generic assessment. Any general criteria of scientific adequacy that might be used to distinguish science from pseudoscience are either too abstract on their own to decide what is scientific or not, or they are contentious. This is not to deny that astrology, so-called creation science, and explicitly racialist sociobiology are clearly quackery or disguised ideology; it is merely to point out that these judgments must be supported case by case, based on specific empirical knowledge.

Institutions can matter: Science has to be studied as it actually works and that requires investigating much more than a rarified logic of explanation and confirmation. Science is surely a social enterprise. It does not follow from this claim that science is a mere social construction, that evidence plays no role in science, or that science has no better epistemic status than any other institution. It is an empirical question whether the institutions, culture, power relationships, and so on of science promote or hinder the pursuit of scientific knowledge (Kitcher 1993). Social scientists, historians, and philosophers of science have indeed produced many illuminating studies of science in practice and treating science scientifically requires asking what role social processes play, but they do not support the more extreme, all-encompassing claims about mere social construction.

Scientific social science: The above discussion of science and pseudoscience should make it obvious that questions about the genuine scientific status of all—or some particular—social science are sensible only if (1) they are posed as questions about specific bodies of social research and (2) they are approached as concrete inquiries into the evidential and explanatory success of that body of work. Assessing scientific standing is continuous with the practice of science itself.

This means that providing all-purpose arguments about what the social sciences can or cannot do on broad conceptual grounds is misguided. The same holds for judging the social sciences by comparison with positivist misunderstandings of physics.

A fair amount of past philosophy of social science was this kind of unfortunate project. For example, Charles Taylor (1971) argued in a widely cited article that the "human" sciences were fundamentally different from the other sciences because explaining human behavior requires understanding meanings and therefore the human sciences cannot provide the kind of "brute" data (Taylor's word) that the natural sciences provide.

There are two clear problems with arguments like this. First, they make blanket claims about the social sciences that are implausible. Lots of social research is not about individual beliefs, interpretations, symbols, and so on. Instead it is about macrolevel or institutional processes. So organizational ecology studies the competitive environment determining the differential survival of organizations (Hannan and Freeman 1989). Individual beliefs and interpretations are not part of the story. There is an implicit individualism in arguments like Taylor's.

Secondly, Taylor's argument has an implict positivist understanding of the natural sciences, which is ironic given that Taylor would certainly not think of himself as holding such views. Data in the natural sciences are acquired and interpreted based on a host of background assumptions and are not "brute." Understanding meanings—and this term hides a host of different things—certainly requires background knowledge, but the question for the social sciences is the same as for the natural sciences: What knowledge is assumed and what is its quality? This general point has been argued by Follesdol (1979), Kincaid (1996), and Mantzavinos (2005). In a way Daniel Dennett's entire project argues something similar. Good social science is aware of the problem that meanings bring and tries to deal with them. For example, careful experimental work in the social sciences goes to great pains to control for subjects' understanding. There are many ways such problems show up in the social sciences and no doubt some social science handles them badly. But it is a case-by-case empirical issue, not a deep conceptual truth about the nature of the human.

Views like Taylor's are a denial of an important—and correct, in my view—doctrine about the social sciences that is a form of naturalism (Kincaid 1996). Human social organization and behavior is part of the natural order and thus amenable to scientific study. No doubt human social behavior raises its own set of difficulties calling for methods not found in physics, for example. But the methods of the natural sciences differ greatly across the sciences as well. Geology, cosmology, and evolutionary biology are much less experimental than other natural sciences, but basic scientific virtues such as ruling out competing explanations are embodied in their practices. Naturalism says that those virtues are possible and necessary in the social sciences as well.

These are the guiding philosophical ideas behind the chapters in this volume. The goal has been to promote work in philosophy of social science that parallels the good work our colleagues in philosophy of biology have produced—work that engages with the science and its ongoing controversies. Plenty of philosophical issues arise but largely in the context of problems in contemporary social research. Given the latter interest, it is not surprising that contemporary developments in social science also strongly influence the chapters included. I want to next discuss some of those developments and in the process survey the issues raised by the various chapters.

1.2. Overview of the Issues

There has been a renewed interest in causality and causal complexity among social scientists that has interacted with other developments in methodology. It is arguable that much social science from the 1950s through the 1970s was suspicious of making causal claims about the social world (Hoover 2004). This suspicion goes back to Hume through Pearson, whose causal skepticism was part of the trimmings of the new statistical methods he helped develop that have been central to much social science. However, social scientists have deep interests in policy and political issues, and thinking about those things requires causal notions. So causal interest never really went away. Some social scientists—primarily economists—started trying to determine the conditions under which regressions could be interpreted causally in the 1950s, and there were further forays later. However, in the last fifteen years the tools for explicit causal modeling have expanded and increased in rigor with groundbreaking contributions from computer science (Pearl 2000) and philosophers of science (Glymour et. al 1987). Explicit causal models are now much more common in the social sciences in part due to these developments. At the same time, philosophers of science took increasing interest in nonreductive accounts of causation and the methods they entail (Cartwright 1989, 2009 and Woodward 2005).

Several other factors also contributed to renewed interest in and confidence about making causal judgments. Movements in sociology have emphasized the importance of mechanisms (Hedström and Swedberg 1998) and mechanisms are naturally explicated by causal notions. A need for such mechanisms was also motivated by the widespread expansion of rational choice game theory and then evolutionary game theory (and related modeling techniques) in social sciences outside economics. Applied game theory provides possible mechanisms for stable macropatterns, raising suspicions of macropatterns without a mechanism.

A third trend that has moved causal thinking to the fore is increasing statistical sophistication in the social sciences, made possible in part by increased computing power. Part of that sophistication appeared in the explicit causal modeling mentioned above, which moved in tandem with application of Bayesian notions in the social sciences. Another source of sophistication that led to more explicit causal thinking was the introduction of large-scale randomized trials into the social sciences and the development of statistical methods such as instrumental variables and potential outcomes analysis (Dufflo, Glennerster, and Kremer 2008, Angrist and Pischke 2008). These methods hope to indentify causes explicitly.

Parallel to increased interest in causality was an increased interest in complex causality. Complex causality is used in various ways, but some standard notions are threshholds, conjuctive causes, and necessary causes. The basic claim is that in the social world, the causes are not thought of as a set of independently acting sufficient causes that operate everywhere and are everywhere the same. These recognitions were embodied in innovative and nontraditional methods for dealing with constellations of causes, using Boolean algebra and fuzzy set theory (Ragin 1987), for example. Anthroplogists had always argued that social causality was complex and contextual, but now sociologists and political scientists were saying the same thing, using new tools to look at their subject matters.

Thus the chapters in part I take up a variety of issues about causality in the social sciences. Petri Ylikoski and I are both concerned with unpacking the claims that social science needs causal mechanisms. Ylikoski argues that on one of the best conceived pictures of mechanisms—that outlined by philosophers of biology—mechanisms in the social sciences argue against various forms of individualism. Mechanisms may certainly make heavy use of agents' perceptions, intentions, and actions. Yet nothing about a proper understanding of mechanisms makes explanations in terms of individuals the full story or the fundamental story. Rather, mechanism-based explanation is largely achieved through interfield accounts from multiple disciplines linking macro and micro in reciprocal ways. It is individual behavior acting in the preexisting institutional and social context that is important. This theme is repeated in part III, Norms, Culture, and the Social-Psychological, in the chapter by David Henderson on norms and by Don Ross on the origins of social intelligence. Both argue that such context is essential for successful explanations to take into account the institutional and cultural factors.

David Waldner continues the discussion of mechanisms by looking at the currently popular idea in the social sciences that process tracing is an important evidential and explanatory strategy, and ties it to a particular understanding of mechanisms. He notes that there is a clear distinction between wanting mechanisms for explanation as opposed to wanting them to provide evidence. Waldner argues that the most interesting understanding of process tracing comes from identifying the mechanisms that underlie established causal relations (what I call vertical mechanisms). Identifying intervening causes between established causal relations (horizontal mechanisms) has value, but it does not explain why causal relations hold. Mechanisms that do so provide explanatory added value and they are not variables as traditionally conceived (they cannot be manipulated independently of the causal relations they bring about), but are invariants—they generate the correlations and causal relations that are observed. Mechanisms in this sense can be individual actions, institutional constraints, and so on and combinations thereof.

On the evidential side, the methods associated with process tracing claim to be different than standard statistical methods. Waldner agrees. Yet he argues persuasively that these alternative methods at present are quite informal and in need of further clarification to establish their reliability. In terms of the philosophies of science sketched earlier, advocates of process tracing realize that social science evidence is not reducible to simple, more or less a priori rules. Yet that does not mean that anything goes, and defending and articulating the reasoning behind process tracing is an important and underdeveloped project essential to advancement in the social sciences.

Julian Reiss ties into Waldner's discussion of process tracing by giving clear conditions and usages for counterfactual claims in the social sciences. He points out that process tracing does not give us information about the actual difference a potential cause makes (which is Robert Northcott's main concern). Counterfactuals can help tell us about such differences. Furthermore, analyzing counterfactuals requires explicit causal models, and developing these can help avoid various biases that often operate when no such model is present (I make a parallel point).

I also point out that the notion of a mechanism can mean multiple different things, that mechanisms can be wanted for different things—for example, for confirmation of causal claims versus for providing causal claims of sufficient explanatory depth—and that the resulting variety of different claims about mechanism need not all fall or stand together. Using the directed acyclic graph (DAG) framework, I argue that there are some specific situations where mechanisms are needed to avoid bias and confounding. Standard regression analysis in the social sciences often misses these problems because they work without explicit causal models. These arguments are about mechanisms in general and give no support to the idea that the mechanisms *must* be given in terms of individuals.

The DAG framework suffers in situations where the causal effect of one factor depends on the value of another. I argue that the DAG formalism has no natural way to represent this and other complex causes such as necessary causes. In part II, Evidence, Stephen Morgan and Christopher Winship present an interesting, novel, and empirically well motivated route for handling a specific subset of interactions in DAGs motivated by the literature on education and outcomes that will be an important contribution to the literature and builds on their previous substantial work on causal modeling in the social sciences (obviously, the evidence and causation chapters overlap). Their results certainly provide another concrete sense of needing mechanisms.

The causal complexity discussed in my chapter and by Morgan and Winship refers to situations where it is unrealistic to think that a particular type of effect is caused by a list of individual causes, each having an independent measurable sufficient partial effect on the outcome. Further complications involved in this picture of social causation are investigated by Northcott and by David Byrne and Emma Uprichard. Northcott's concern is finding coherent accounts of causal effect size in the existing (mostly regression based) literature. To put the moral in brief, regression coefficients are not generally good measures of effect size or causal strength and even when they are, they depend strongly upon already having good evidence about the causal relations or structure in play, a point emphasized by Northcott as well as myself. Byrne and Uprichard discuss varieties of causal complexity—in cases where it is not realistic to think that the string of independent causes model applies—and methods for dealing with them. In particular, they focus on the qualitative comparative analysis framework of Ragin using Boolean logic and fuzzy set theory that promises to go beyond standard correlation statistics when dealing with complex causes. That framework deserves more discussion than space allowed for in this volume—it deals with complex causation in a way that philosophers would naturally understand and it has novel methodological tools that are becoming increasingly popular.

Gary Goertz picks up on the limitations of standard statistical methods for confirming causal claims. His chapter is full of rich, interesting examples of social science causal-descriptive generalizations that are well established, despite the common mantra that none such exist. He makes an important point that seems obvious once it is understood but is not widely grasped: A set-theoretical claim of all As are Bs can be consistent with zero correlation in statistical senses. In terms of the philosophy of science sketched at the beginning, statistical reasoning relies on a formal logic of inference that does not handle all relevant complexities.

A deeper, more philosophical issue lying behind work on causality in the social sciences concerns understanding the probabilities they support. While it is possible to interpret probabilities in social research as resulting from measurement imprecision or from unmeasured variables, these are not entirely satisfactory accounts. It seems that we end up with probabilistic causes even when our measurements are quite reliable. Second, why should unmeasured causes produce the kinds of stable frequencies that we see in the social realm? Marshall Abrams provided a sophisticated answer in terms of a novel account of what he calls mechanistic probability—stable frequencies produced from underlying causal processes with specific structure. Such structures exist in nature—a roulette wheel is a paradigm instance—and there is good reason to think that in the social realm there are social equivalents of roulette wheels.

Part II of the volume contains chapters about evidence. Of course, chapters in part I are also concerned with evidence, and explanation issues show up in part II. However, there is a decisive shift in emphasis in the chapters of the two parts.

Fred Chernoff surveys the history up to the present of the Duhem's underdetermination thesis. He notes that it is not nearly as radical as Quine's, which I argued earlier was excessive and ignored the variety of techniques that scientists can use to triangulate on where to place blame when hypotheses do not match the data. Duhem's concern was to deny that simply by the use of formal deductive logic, one could determine with certainty whether a hypothesis was confirmed or not. In short, he was a precursor of the postpositivist philosophy of science sketched earlier that rejects the logic of science model. Assessing the evidence depended upon the good common sense of the relevant scientific community.

Chernoff also discusses the relevance of Duhem's view that there may be multiple ways to measure or operationalize aspects of theories, and in that sense which measure is used is conventional. Duhem did not think that this made the choice arbitrary—the good common sense of the scientific community was again needed—but that adopting a common measuring procedure was crucial for scientific progress. Chernoff provides a detailed case study of two important areas in international relations—the democratic peace hypothesis and balance of power theories—showing how in the former common measures promoted significant scientific progress, and the lack of them in the latter undermines its empirical qualifications.

Andrew Gelman and Cosma Rohilla Shalizi discuss the use of Bayesian methods in social science testing based on their considerable combined experience. However, their take on Bayesian methods is quite different than the usual subjective Bayesians versus objective frequentist debate. That debate is often framed as being about which of these views is the true logic of science, and thus based on a false presupposition from the postpositivist point of view. Gelman and Shalizi don't see much value in the exercise of starting with subjective priors and updating them to a new posterior distribution. However, they argue that Bayesian methods are quite useful when it comes to model checking in the social sciences. Model checking as they mean it is a paradigm instance of the kind of piecemeal triangulation that radical holists miss.

Aviezer Tucker also uses Bayesian ideas in his discussion of the relation between the social sciences and history. He argues that history is not applied social science, and social science is generally not history. History is about inferring to common cause token events in the past using background theories of information transfer applied to currently available traces in the form of such evidence as documents. Social science is about relating types—variables—by quite different, often statistical, methods. Bayesian ideas come into play in two ways. He argues that inferring to a past token event as a common cause of multiple present information traces is a matter of the likelihood of the common cause hypothesis versus its competitor. That is not a fully Bayesian framework, because it does not involve priors. However, Tucker argues that social science results can tell historians what possible past tokens are initially plausible as common causes. Inferring who wrote the Bible can be informed by the finding that writing only arises in the presence of a centralized bureaucratic state, and thus that books of the Old Testament cannot be contemporary to the events they described. In that sense the social sciences can provide priors. However, priors in this sense are just relevant background information—in other words, good scientific common sense.

Nancy Cartwright's chapter on randomized controlled trials (RCTs) as evidence for potential policy effectiveness echoes the general theme of part II that evaluating evidence in practice is a complex and fallible affair that rules of scientific logic do not capture. RCTs are treated by the medical profession and increasingly by social scientists—they are all the rage in development economics, for example—as the gold standard. That phrase is widely used without clear explanation, but it generally means either that RCTs are thought to be near conclusive proof, the only real proof, or by far and away the best proof. In short, their logic guarantees reliable outcomes, another of the hopes for a logic of science. Cartwright argues convincingly and in detail that RCTs can be quite unreliable as guides to policy effectiveness.

Morgan and Winship take up in much greater detail the issues raised by interaction effects and heterogeneity for DAG analyses that I raise in my chapter. They provide an explicit framework for incorporating such complications into DAGs. Their basic approach to the possible errors caused by interaction and heterogeneity is to model them. Like Gelman and Shalizi, their concerns are driven by the kinds of problems they see in existing research, which in their case are the causes of educational attainment. Formal methods like DAGs are useful, but their usefulness has to be evaluated according to the kinds of causal complexity faced by practicing researchers and adapted accordingly. They note that the formalism of DAG models can be a hindrance to recognizing causal complexities.

Ken Kollman's chapter continues the emphasis on the complexities of evidence, focusing on the burgeoning field of computational models of social phenomena. In one way his topic is a classical one, especially in philosophy of economics, about the status of abstract and idealized models. Kollman notes what modelers often say in their defense—namely, that models provide insight. However, he goes a step further and realizes that appeals to insight are not enough (it could be a warm and fuzzy feeling only, though this is my formulation, not his). Kollman gives several other, more concrete reasons such models may be reasonable. It is possible to generate simulated data with computational models and then compare the patterns in the data with real empirical patterns in analogous social data. So empirical testing is possible, though Kollman cautiously notes that there are still issues about how strong the analogy is. Computational models also have explanatory virtues: They instantiate the causal mechanical ideals advocated in the chapters in part I. This means they can represent dynamics, something that rational choice game theory, for example, cannot. He also argues that they provide ways to model micro and macro social phenomena, in line with Ylikoski and Waldner's idea that mechanism-based explanation defuses individualism/holism debates.

The chapters in part III deal with an intersecting set of topics concerning culture, norms, and the explanation of sociality. Here issues of explanation (macro and micro, for example), evidence, and more philosophical issues concerning how to understand key concepts are intertwined. Most chapters ask the question: How do explanations in terms of norms, culture, and related concepts relate to psychological explanations? To what extent are the latter sufficient? Necessary? What is the basis of human sociality? Human nature or social organization or some mix of the two? And if the latter, how does that work?

Mark Risjord provides a history up to the present of the concept of culture in anthropology, where the concept is most used. That history has been a running conflict between treating culture as a trait of individuals—a form of methodological individualism—and as something superseding individuals and sometimes indeed as controlling them. The most plausible view, according to Risjord (echoing the approach emphasized by Ylikoski) is to see that debate as dissipated by a more interactive view where neither the individualist or holist view is on the table. Though that is a common theme throughout the volume, there is obviously more work to do in fleshing out that claim. My guess is that there are multiple, domain-specific ways of doing so, and I would not claim that this volume is anything like the final word on the issue.

Henderson takes on clarifying norms, a concept widespread throughout the social sciences, though it is generally not carefully explicated. Sometimes norms are only behavioral regularities. Henderson argues convincingly that in this guise they are not particularly explanatory. His main focus is on norms as knowing (and having attitudes about) a rule, following some of the most sophisticated recent analyses. Henderson argues that rules cannot be seen as entirely a psychological phenomenon, because payoffs and differences in social status and power are part of the explanation. However, there are important questions, largely unexplored in the literature, about the psychological basis and explanation of knowing rules. To what extent can cognitive science accounts be integrated with sociological, economic, and anthropological accounts? Like Ylikoski in chapter 2, Henderson thinks that an interfield account is called for.

The evolutionary program in social science is the subject of Francesco Guala's and Tim Lewens's chapters. Guala focuses particularly on the debate over whether cooperation and sociality in humans requires strong reciprocity—roughly, the willingness to perform costly sanctions to enforce norms—or can be simply explained in terms of self-interest. This empirical issue is important for policy decisions, since if humans are not generally capable of strong reciprocity, policies that assume they are will lead to bad outcomes. Lewens provides an overview of objections to theories of cultural evolution. He delineates the relation between sociobiology and other kinds of evolutionary accounts and between meme-based versions and population level learning accounts. Levens give us a balanced account that argues that not all problems raised in the literature against evolutionary models are decisive, and yet is wary of attempts to push further than we can go.

Ross looks at the interactive origins of human intelligence and sociality, specifically at the thesis that human intelligence in evolutionary history resulted from the need to meet the needs of social interactions. He surveys neurobiological and other evidence that suggests primates in general have natural dispositions to cooperate. So human intelligence seems unlikely to be the result simply of the need for social coordination. Instead, Ross suggests that that when hominid groups developed specialization and trading, greater demands arose to deal with these new forms coordination. Complex socialized selves were needed to play the more complex games that exchange and specialization involves.

Ron Mallon and Daniel Kelly examine the status of race as a social science concept. The biological notion of race seems quite unfounded, so how has it been a useful concept in the social sciences—or has it been? They deny that race is fully explained as a social role and argue that there is important empirical evidence suggesting that there are strong psychological underpinnings behind our tendencies to categorize people in terms of race. This is in keeping with the theme of many chapters that macro and micro accounts need to be involved and integrated.

The chapters in part IV focus on issues in the sociology of knowledge. Earlier chapters had already informally considered some sociological and rhetorical aspects of social science research. As argued earlier in the chapter, information about the sociological factors driving research can be useful information in assessing the scientific standing of various lines of research.

Amy Mazur discusses feminist social science research, especially feminist comparative politics (FCP), her prime area of interest. The feminist research she advocates and discusses aims to contribute to accumulation of knowledge through empirical research, and she carefully distinguishes this from extreme constructivist views about science that some feminists have espoused. The feminist research she advocates does proceed, however, with an awareness of and interest in gender issues and a recognition of how gender biases can infect standard social science research.

She details the empirical success of feminist comparative politics. Mazur describes the social organization of the FCP community and its interaction with elements of national governments that have made it a success. However, she notes that mainstream comparative politics has largely ignored these achievements and argues that gender biases continue to plague the mainstream, which is still largely comprised of male researchers.

Allan Horwitz applies the sociology of knowledge approach to mental illness. He rejects the idea that the sociology of mental illness classification and organizational embeddedness shows that mental illness is a pure social construct (just as Mazur rejects radical constructivist feminist views about science). He also thinks that saying that all mental illness is a matter of looping kinds—interaction between individual traits and the effect on the individual classified as having some mental disorder—as Hacking sometimes suggests is too crude a formulation that glosses important differences. Looping seemingly plays a much bigger role in ADHD than it does in schizophrenia. Horwitz believes that there can be neurobiologically based mental malfunctions that constitute mental illness. Looking at the social and institutional processes involved in the classification and treatment of behavior of mental disorders can be quite helpful in assessing which current practices have a grounded basis and which ones exist largely due to the sociology of the psychiatric profession and the classification process.

The final chapters of the volume comprising part V focus on normative issues that have important ties to social science research and philosophy of science issues. James Woodward uses the kind of work on reciprocity in cooperative behavior discussed by Guala and Ross to ask what implications it may have for political philosophy. Daniel Hausman discusses the difficulties in evaluating health outcomes in terms of the preferences of patients and concludes that evaluation often relies on messy ad hoc processes. Anna Alexandrova asks if social science research on well-being actually gets at well-being (something its critics wonder about). She argues that philosophical accounts of well-being are of minimal help, and in practice the different sciences that study well-being use different, local notions relevant to the context without compromising their results. This is in keeping with the post-positivist moral drawn at the beginning that science often does not work with concepts definable in terms of necessary and sufficient conditions.

NOTES

Parts of this introduction are taken from Ross, D., and Kincaid, H. "The New Philosophy of Economics," in H. Kincaid and D. Ross, eds., *Oxford Handbook of the Philosophy of Economics* (Oxford: Oxford University Press, 2009), 3–35.

1. Another distinct difference is over the role of values. Since I have pursued this at length elsewhere (Kincaid 2007), I am not going to do so systematically here. There are many different ways values can be involved with different consequences. The short answer

is that science is a complex set of practice and that values can cause bias in some cases but not in others. For example, Mazur's chapter shows how values can both lead to better science and to bad science as does Horwitz's chapter on mental illness.

REFERENCES

Angrist, J., and J-S. Piscke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton, NJ: Princeton University Press.

Cartwright, N. 1980. "Causality in a World of Instrumental Laws." *PSA: Proceedings of the Biennial Meeting of the Philosophy of Science Association* 2: 38–48.

Cartwright, N. 1983. *How The Laws of Physics Lie*. Oxford: Oxford University Press.

Cartwright, N. 1989. *Nature's Capacities and Their Measurement*. Oxford: Oxford University Press.

Cartwright, N. 2007. *Hunting Causes and Using Them*. Cambridge: Cambridge University Press.

Duflo, E., R. Glennerster, and M. Kremer. 2008. "Using Randomization in Development Economics Research: A Toolkit." In *Handbook of Development Economics*, T. P. Schultz, and J. Strauss, eds., 3895–3962 Amsterdam: Elsevier.

Follesdol, D. 1979. "Hermenutics and the Hypothetical-Deductive Method." *Dialectica* 33: 319–36.

Glymour, C. 1980. *Theory and Evidence*. Princeton, NJ: Princeton University Press.

Glymour, C., R. Scheines, P. Spirtes, and K. Kelly. 1987. *Discovering Causal Structure: Artificial Intelligence, Philosophy of Science, and Statistical Modeling*. New York: Academic Press.

Hannan, M., and J. Freeman. 1989. *Organizational Ecology*. Cambridge, MA: Harvard University Press.

Hedström, P., and R. Swedberg. 1998. *Social Mechanisms: An Analytical Approach to Social Theory*. Cambridge: Cambridge University Press.

Hoover, K. 2004. "Lost Causes." *Journal of the History of Economic Thought* 26: 129–44.

Kincaid, Harold. 1996. *Philosophical Foundations of the Social Sciences*. Cambridge: Cambridge University Press.

Kitcher, P. 1993. *The Advancement of Science*. Oxford: Oxford University Press.

Ladyman and Ross. 2007. *Everything Must Go*. Oxford: Oxford University Press.

Mantzavinos, C. 2005. *Naturalistic Hermeneutics*. Cambridge: Cambridge University Press.

Moss, L. 2004. *What Genes Can't Do*. Cambridge, MA: The MIT Press.

Pearl, Judea. 2000. *Causality: Models, Reasoning, and Evidence*. Cambridge: Cambridge University Press.

Ragin, Charles. 1987. *The Comparative Method: Moving Beyond Qualitative and Quantitative Methods?* Berkeley: University of California Press.

Taylor, Charles. 1971. "Interpretation and the Sciences of Man." *Review of Metaphysics* 25: 3–51.

Wilson, M. 2007. *Wandering Significance*. Oxford: Oxford University Press.

Woodward, J. 2005. *Making Things Happen: A Theory of Causal Explanation*. Oxford: Oxford University Press.

This page intentionally left blank

PART I

Mechanisms, Explanation, and Causation

This page intentionally left blank

CHAPTER 2

MICRO, MACRO, AND MECHANISMS

PETRI YLIKOSKI

2.1. Introduction

This chapter takes a fresh look at micro-macro relations in the social sciences from the point of view of the mechanistic account of explanation. Traditionally, micro-macro issues have been assimilated to the problem of methodological individualism (Udéhn 2001, Zahle 2006). It is not my intention to resurrect this notoriously unfruitful controversy. On the contrary, the main thrust of this chapter is to show that the cul-de-sac of that debate can be avoided if we give up some of its presuppositions. The debate about methodological individualism is based on assumptions about explanation, and once we change those assumptions, the whole argumentative landscape changes.

The idea that social scientific explanations are based on causal mechanisms rather than covering laws has become increasingly popular over the last twenty years or so (Hedström and Ylikoski 2010). Interestingly, a similar mechanistic turn has occurred also in the philosophies of biology and psychology (Wright and Bechtel 2007). Until recently, the connections between these two emerging traditions for thinking about mechanisms have been rare. The aim of this chapter is to employ ideas developed by philosophers of biology to address some issues that the advocates of mechanisms in the social sciences have not yet systematically addressed. I argue that ideas about levels of explanation and reductive research strategies, which were originally developed in the context of cell biology and neuroscience, can be fruitfully adapted to the social sciences. They can both strengthen the case for mechanism-based explanations in the social sciences and bring the philosophy of social science debates closer to social scientific practice.

The chapter is structured as follows. In the first section, I will take a look at recent work on mechanism-based explanation. While I suggest that the mechanistic account of explanation presupposes some more fundamental ideas about explanatory relevance and causation, I also argue that it provides a fruitful tool for thinking about micro-macro relations in the social sciences. In the second section, I will criticize a common philosophical way of formulating the micro-macro issue and provide my own characterization that is not dependent on the assumption that there is a unique or comprehensive micro level. The third section introduces the distinction between causal and constitutive explanation, and argues that this distinction helps to make sense of the call for microfoundations in the social sciences. The final section will take on a doctrine that I call intentional fundamentalism, and it challenges the idea that intentional explanations have a privileged position in the social sciences.

2.2. Mechanism-based Explanation

The idea of mechanism-based explanation has been developed independently among social scientists (Harré 1970; Elster 1989, 2007; Little 1991; Hedström and Swedberg 1998; Hedström 2005; for a review see Hedström and Ylikoski 2010) and philosophers of biology (Bechtel 2006, 2008; Craver 2007; Darden 2006; Wimsatt 2007). In the social sciences, the idea of causal mechanism has been used mainly as a tool for methodological criticism, while in the philosophy of biology the motivation has been that of finding a descriptively adequate account of biological explanation. Despite these separate origins and motivations, both traditions are clearly building on similar ideas about scientific explanation. For example, both share the same dissatisfaction with the covering law account of explanation (Hedström 2005; Craver 2007).

There is no consensus on the right definition of a causal mechanism. Although some theorists find such a situation frustrating, I do not think this constitutes a real problem. The entities and processes studied by different sciences are quite heterogeneous, and it is probably impossible to propose a mechanism definition that would be both informative and cover all the prominent examples of mechanisms. Some disciplines, such as cell biology (Bechtel 2006) and the neurosciences (Craver 2007), study highly integrated systems, whereas others, such as evolutionary biology and the social sciences, study more dispersed phenomena (Kuorikoski 2009), so it is more plausible to think that informative characterizations of mechanisms are field specific. The task of a philosophical account is to show how these exemplars are related to general ideas about explanation, evidence, and causation, not to engage in verbal sophistry. However, it is possible to give some general characteristics of mechanisms.

First, a mechanism is always a *mechanism for something*; it is identified by the kind of effect or phenomenon it produces. Second, a mechanism is an *irreducibly* *causal notion*. It refers to the entities of a causal process that produces the effect of interest. Third, a mechanism has a *structure*. When a mechanism-based explanation opens the black box, it makes visible how the participating entities and their properties, activities, and relations produce the effect of interest. The focus on mechanisms breaks up the original explanation-seeking why-question into a series of smaller questions about the causal process: What are the participating entities, and what are their relevant properties? How are the interactions of these entities organized (both spatially and temporally)? What factors could prevent or modify the outcome? Finally, there is *a hierarchy of mechanisms*. While a mechanism at one level presupposes or takes for granted the existence of certain entities with characteristic properties and activities, it is expected that there are lower-level mechanisms that will explain them. In other words, the explanations employed by one field always *bottom out* somewhere. However, this fundamental status of certain entities, properties, and activities for a given mechanism is only relative, as they are legitimate targets of mechanistic explanation in another field. Of course, this chain of explanations ends somewhere—there are no mechanism-based explanations for fundamental (physical) processes (Hedström and Ylikoski 2010).

Although the mechanism-based account is often presented simply as an idea about scientific explanation, the notion of mechanism is associated with a wider set of ideas about scientific knowledge. For example, there are ideas about the justification of causal claims, the heuristics of causal discovery, the presentation of explanatory information, and the organization of scientific knowledge (Ylikoski 2011). There is no doubt that these not yet clearly articulated ideas partly explain the appeal of the approach. For example, as I will show later in this chapter, claims about the explanatory role of mechanisms are often confused with claims about their relevance to the justification of causal claims (see also Kincaid, this volume).

While I think all the above ideas are important advances in understanding explanatory reasoning in science, it is not necessary to assume that the notion of mechanism is the ultimate solution to all problems in the theory of explanation. On the contrary, the mechanistic theory presupposes accounts of explanatory relevance, causation, and the nature of generalizations that provide the basis for mechanisms. The notion of mechanism should not be treated like a black box. I have argued elsewhere (Hedström and Ylikoski 2010; Ylikoski 2011) that if the mechanistic ideas are combined with the theory of explanation developed by James Woodward (2002, 2003), we can get quite far in solving these problems. While for the present purposes we do not have to consider in detail the relation between mechanisms and generalizations, some comments on explanatory relevance are in order as later arguments depend on it.

A mechanism-based explanation describes the causal process selectively. It does not aim at an exhaustive account of all details but seeks to capture the crucial elements of the process by abstracting away the irrelevant details. The relevance of entities, their properties, and their interactions is determined by their ability to make a relevant difference to the outcome of interest. If the presence of an entity or of changes in its properties or activities truly does not make any difference to the effect to be explained, it can be ignored. This counterfactual criterion of relevance implies that mechanism-based explanations involve counterfactual reasoning about possible changes and their consequences (Ylikoski 2011). A natural way to understand these causal counterfactuals is to understand them as claims about the consequences of ideal causal interventions (Woodward 2003, 2008). The causal counterfactual tells us what would have happened to the effect if the cause had been subject to a surgical intervention that would not have affected anything else in the causal configuration. An advantage of the interventionist account of causation is that it allows talking about causal dependencies in every context where the notion of intervention makes sense. Unlike some other theories of causation, such as various process theories, it is level-blind and applicable to special sciences such as cell biology or sociology.

2.2.1. Mechanisms and Reductive Explanation

One of the distinctive features of the mechanistic approach to explanation is that it reorients the issues related to reductionism and reductive explanation. In one sense the mechanistic way of thinking is thoroughly reductionist: It attempts to explain activities of mechanisms in terms of their component parts and their activities, and then subjects the component mechanisms to the same treatment. In this sense, the reductive research strategy has probably been the single most effective research strategy in the history of modern science. However, there is another sense in which mechanism-based explanations are clearly nonreductionist: Although they do refer to the micro level, they do not replace or eliminate the higher-level facts nor the explanations citing them. Rather than serving to reduce one level to another, mechanisms bridge levels (Darden 2006; Craver 2007; Wright and Bechtel 2007; Richardson 2007; McCauley 2007; Wimsatt 2007).

The mechanistic account of reductive explanation differs significantly from the traditional philosophical accounts of intertheoretical reduction that conceive reduction as a derivation of one theory from another (Richardson 2007; McCauley 2007). The mechanical account of reductive explanation does not start with a strongly idealized picture of a discipline-wide theory that contains all knowledge about its level. Nor does it conceive reduction as a deductive relation between such theories (or their corrected versions). Rather, reductive mechanistic explanations are constructed piecemeal with a focus on particular explanatory targets. While there is an assumption that everything is mechanistically explainable and a presumption that ultimately all mechanistic accounts are mutually compatible, there is no overarching effort to combine them into one grand theory that would cover all the phenomena that the scientific field studies. Also, contrary to the traditional accounts that conceive reduction as elimination or replacement, the mechanisms are inherently multilevel. The components and their operations occur and are investigated at one level, whereas the mechanism itself and its activities occur and are investigated at a higher level. In this sense accounts of mechanisms often have the character of interfield theory (Darden 2006). This makes it difficult to characterize the reductive understanding provided by mechanical explanations as deductive relations between independent theories.

The mechanistic stance also gives reasons for rethinking the notion of levels. According to the traditional layer-cake conception, there is a neat hierarchical layering of entities into levels across phenomena, and the scientific disciplines (e.g., physics, chemistry, biology, psychology, sociology) are distinguished from each other by the level of the phenomena that they are studying (see Oppenheim and Putnam 1958). From the mechanistic point of view, this way of thinking unnecessarily drives together levels of nature and science, and misleadingly suggests that the levels are both comprehensive and the same independently of the investigative context (Craver 2007). The actual scientific disciplines do not match neatly with the metaphysical picture of levels of organization or reality. And while there are many problems in a serious characterization of the metaphysical picture of levels, there do not seem to be any particularly good reasons to accept such a metaphysical constraint for an account of scientific explanation.

The notion of the levels of mechanism plays an important role in the mechanistic account but is free from many of the traditional assumptions about levels. The levels of mechanisms are perspectival in the sense that the levels are dependent on the explanatory target. Macro-level facts are explained by appealing to micro-level processes, entities, and relations, but these items belong to the micro level just because they are required for the full explanation of the macro fact, not because they belong to some predetermined micro level. Whatever is needed for explaining the macro fact is regarded as belonging to the same level. However, there is no guarantee that these components would always be at the same level in all possible explanatory contexts. Nor it is obvious that the micro-level entities and processes that account for these components would be in any simple sense from the same level. For every hierarchy of mechanisms, there is a clear hierarchy of levels of mechanisms, but these levels are local. There is no reason to assume that separate hierarchies of mechanism levels would together produce the neatly delineated and comprehensive levels of nature assumed in the traditional layer-cake model (Craver 2007).

These views have a number of interesting consequences for traditional ways of thinking about reductive explanation and the explanatory role of microfoundations in the social sciences. For example, once we give up the outdated deductive model of theory reduction, many of the traditional fixations of the methodological individualism debate simply become meaningless. For example, there is no need to provide individualistically acceptable redefinitions of macro-social notions because the explanation of macro facts is no longer conceived as a logical derivation. Similarly, the search for any bridge laws between theories becomes pointless. This has the consequence that the key anti-reductionist argument about multiple realization loses much of its significance. From the point of view of mechanistic explanation, multiple realization is simply an interesting empirical observation that does not pose any serious threat to the possibility of explaining macro properties in terms of micro properties and relations. Just as the sciences have learned to live with the fact of alternative causes, they can learn to live with the phenomenon of multiple realization.

The advocates of the mechanism-based approach in the social sciences have noticed some of these consequences. For example, they have largely given up the old ideas about reductive explanation and have instead emphasized the importance of microfoundations (Elster 1989; Little 1991). However, I do not think that all the implications of the mechanistic perspective have been taken into account in the philosophy of social sciences. This is visible, for example, in the fact that quite often the mechanistic approach is associated with methodological individualism (Elster 1989). Similarly, much of the debate about micro-macro relations is still focused on arguments that are based on a premechanistic understanding of reductive explanation (Sawyer 2005; Zahle 2006).

The aim of this chapter is to sketch what a consistently mechanistic way to think about micro-macro relations would look like and to show that some of the key presuppositions of the traditional debate about methodological individualism should be given up. One of these is the assumption of a comprehensive, unique, and privileged individual level. The notion of *comprehensiveness* refers to the idea that there is a consistent and well-defined individual level that is sufficient to cover all social phenomena and that would serve as a reduction basis for all nonindividual social notions. *Uniqueness* refers to the assumption that in all social explanations, the micro level would always be the same level, for example, the level of intentional rational action. Finally, the notion of *privileged* refers to the presumption that explanations in terms of this special level have some special explanatory qualities that set them apart from explanations from other levels. In the following, I will challenge all three assumptions and argue that once they are given up, we can approach the micro-macro issues in the social sciences in a more clear-headed manner.

2.3. Rethinking the Macro

A popular argumentative strategy among anti-individualists has been to borrow ideas from the philosophy of mind. They are inspired by the arguments for nonreductive materialism, so they build their argument based on an analogy with the mind-brain relation. Given that these arguments are not very mind specific—it is a general practice just to talk about M- and P-predicates—their appeal is understandable. The ideas of supervenience and multiple realization seem to provide a neat way to argue against reductionism, at least if one accepts the traditional idea of derivational reduction. While there are reasons to suspect that the notion of supervenience is less illuminating than is often assumed (Horgan 1993; Kim 1993) and that the traditional view of reduction does not completely collapse under multiple realization (Kim 1998), we can set these issues aside as their relevance presupposes a premechanistic account of reductive explanation. Here I want to focus on the mind-brain analogy as I think it is misleading.

The mind-brain analogy is inappropriate because it mischaracterizes the nature of the social scientific micro-macro problem. The central problem in the philosophy of mind is to figure out how the explanations provided by psychological theories that employ mental concepts are related to the accounts of the brain's working provided by the neurosciences. The challenge is to relate two levels of description that are fundamentally talking about the same thing. The (nondualist) antireductionist position does not typically challenge the causal sufficiency of the neural-level facts. The setup is quite different in the social scientific micro-macro debates.

The problem in the social sciences is not that of bridging a comprehensive and exhaustive individual-level understanding of social processes (the analogue to the idealized knowledge of the brain) to a more social or holistic description (that would be analogue to the idealized psychological theories employing the mental vocabulary). It is typical for anti-individualists to challenge the causal sufficiency of individual facts. They often claim that the facts about individuals allowed by the individualist are either not sufficient to account for all social facts or the individualists are cheating by accepting facts that are not properly individualistic. This is because the issue is not really that of relations between two comprehensive (and potentially competing) levels of description, but that of seeing how local facts about individuals and their social interactions are related to large-scale facts about groups, organizations, and societies. So, the relation is really more like the one between the whole brain and its parts than the mind and the brain. While this contrast is useful for highlighting the inappropriateness of the mind-brain analogy, I do not want to develop it further as there are many problems with the organ-society analogy. It is better to skip all the brainy analogies and to take a fresh look at the micro-macro problem as the social scientists face it.

A useful starting point is the observation that macro social facts are typically *supra-individual*: They are attributed to groups, communities, populations, and organizations, but not to individuals. There might be some attributes that apply both to individuals and collectives, but typically macro social properties, relations, and events are such that they are not about individuals.

Another salient feature of many social micro-macro relations is the part-whole relationship. One way or another, the macro social entities *are made of* their constituting parts. Usually this relation of constitution is more than mere mereological aggregation or simple material constitution. First, many social wholes are composed of a heterogeneous set of entities; there are intentional agents, their ideas, and material artifacts. Second, in all interesting examples of social wholes, the *relations* between the components play an important role. (Similarly, often the relations between social wholes and between the social whole and its environment are also important.) However, the important thing is that the part-whole relationship makes it possible to see the micro-macro relation as a question of scale: The difference between micro and macro is the difference between small- and large-scale social phenomena.

I do not propose that we can simply define the micro-macro contrast as an issue of scale. All differences in scale do not constitute a meaningful micro-macro relation, and the heterogeneous nature of macro social facts makes it difficult to characterize the additional requirements for their defining features. However, I do want to suggest that it provides a fruitful way to think about micro-macro relations and an antidote for the tendency to see parallels in the philosophy of mind.

Thinking of the micro-macro issue as an issue of scale makes it possible to conceive of it as being without a unique micro level. Whereas the contrast between "individual" and "social" levels is categorical, the contrast between small and large is relative and allows a continuum of various sizes. Depending on the application, the micro entities could be individuals, families, firms, or groups. This flexibility is in accordance with the way social scientists think. They do not assume that micro is always about one specific set of entities.

Another consequence is that whether an attribute is a macro or micro property depends on what it is contrasted with. A friendship relationship is a macro property from the psychological point of view, but a micro property when considered from the point of view of the social networks within a community. Rather than being set a priori, the contrast between micro and macro depends on one's explanatory interests. For example, international politics and organizational sociology construct the micro-macro contrast quite differently. In the former, states and other organizations are often treated as individuals, whereas in the latter, the organizations and their properties are the macro reality to be explained. Similarly, an economist studying market processes can treat firms and households as the micro level, while for disciplines such as industrial organization and family sociology, they are the macro items that require explanation.

From the point of view of a mechanistic philosophy of science, this flexibility is not surprising. The same dependence of levels on epistemic concerns is also observable in the biological sciences. The cell biologists or neuroscientists do not think in terms of comprehensive or unique micro levels either. The levels of mechanisms found there depend on the explanatory concerns, not on a priori ontological considerations. This is not worrisome for the mechanistic point of view, as the key assumption is that whatever is found at the micro level can always be turned to a macro-level *explanandum* for another set of enquiries.

The social macro properties do not constitute a unified kind, so it makes sense to characterize them with a sample of examples rather than with a general definition. The following classification of typical sociological macro social properties is not intended to be exhaustive of sociology or the social sciences in general. There are many parts of macro social reality that fall between my four categories. However, I hope the four examples can be used to illustrate the applicability of the scale perspective.

2.3.1. Statistical Properties of a Population

A major concern for sociology is the various statistical attributes of populations. Among these are *distributions* and *frequencies*. Sociologists are interested in both distributions of attributes to various kinds of individuals and distributions of individuals with certain attributes to social positions and spatial locations. For example, when they are studying the ethnic segregation of cities, comparing societies in terms of inequality, or describing the social stratification of a society, they are attempting to account for distributions. Another relevant property of distributions are frequencies. Sociologists are interested in typical, rare, dominant, or marginal behaviors, beliefs, or attitudes within a specified population. Similarly, they are interested in ratios of attributes such as unemployment or incarceration within the population. So, when sociologists are studying changes in racial prejudices over time, comparing the level of conformism between communities or tracking the changes in the level of union memberships, they are interested in explaining frequencies.

All these statistical macro social properties are inferred (or estimated) from data about the members of a population. There is no other way to access them. However, it does not make any sense to attribute these properties to individual units. Another important thing about these macro social facts is that the units of these statistics do not have to be individuals; they can as well be families or firms. It is noticeable that statistical macro properties are in no way dependent on the members' beliefs and attitudes about them. The members of the population can have false, or even crazy, beliefs about distributions and frequencies that characterize their own society.

While the statistical properties of populations usually only serve as *explananda* in the social sciences, they do have some legitimate and nonreducible explanatory uses. For example, in the cases of frequency-dependent causation (e.g., cases in which the causal effect of an individual having a certain property depends on the frequency of that property in the population), the statistical facts are the crucial difference makers. Similarly, in many social scientific explanations, the correlations between various variables (for example, wealth, education, taste, and place of residence) play an important role in accounting for individual differences in behavior and attitudes. Both of these cases are quite easily conceived as cases of larger-scale facts influencing smaller-scale phenomena, while other ways to think about levels are not as natural.

2.3.2. Topologies of Social Networks within a Population

Sociologists are also interested in relations and interactions between individuals. When considered together, these relations constitute networks of social relations within the population. A social network can be regarded as a map of all of the relevant ties between the members of a specified population. When sociologists are studying the spread of information within an organization, comparing groups with respect to their level of network clustering or analyzing the brokering opportunities of an individual occupying a structural hole (i.e., a position between two networks that are not otherwise connected), they are examining social networks.

The importance of social networks is increasingly being recognized in the social sciences, and social network analysis is becoming increasingly popular in various social sciences. Social network analysis is based on the observation that networks have many interesting (formal) properties, such as centralization, cohesion, density, and structural cohesion (Scott 2000). While the social network is inferred from knowledge about individual relationships, the properties of the network are prototypical macro properties. It does not make any sense to apply these attributes to individual nodes of the network. Similarly to statistical properties, the units of network analysis are flexible. There is no requirement that the nodes of the network (the members of the population) are persons. They can also be groups, families, organizations, or even states.

The properties of social networks serve both as the *explananda* and the *explanantia* in sociology. As an example of the latter, consider the notion of a structural hole (Burt 1992), which is used to explain the differences in agents' ability to access information and in their opportunities to influence social processes. In these explanations the structure of the network plays an irreducible role, and it is quite natural to think of the social network as a large-scale social phenomenon influencing local interactions between individuals. In contrast, it is very difficult to think about them in terms of social and individual levels. As social networks are attributes of the population, it would be quite a stretch to call social networks individual properties. But if they are macro-level properties, what would be the individual-level properties that could be regarded as their bases? Collections of relevant individual relations, one might suggest, but that would be just a vague way to talk about networks. Things are simpler if one does not have to bother with such questions. A network is simply a more extensive entity that is constituted by more local relations and it can have properties that are not properties of its components.

2.3.3. Communal Properties

By communal properties I refer to social scientific notions that apply to specific communities, but not to isolated individuals. Among these notions are such things as culture, customs, social norms, and so on. For example, cultural differences are primarily between groups, not between individuals. Similarly, social norms and customs are properties of communities—attributing them to solitary individuals does not make sense. Many of these notions do not have precise definitions, and their explanatory uses are often confusing (Turner 1994; Ylikoski 2003), but they do have an important role in the social sciences.

While communal properties are attributed to groups, they are quite straightforwardly based on facts about individuals. Underlying these notions is the idea that the members of a group share certain beliefs, expectations, preferences, and habits. However, it is crucial that the sharing of these individual attributes is not purely accidental: The members have these individual properties because the other members of the group have them. The sharing of these properties is due to continuing interaction. For example, the existence of a social custom presupposes that the novices learn specific expectations and habits when they become members and that the members of the group stick to these expectations and habits because others also do so. Underlying the (relative) unity of a culture are facts about the shared origins of the ideas, values, and practices of the members and their constant interaction with each other. Similarly, the cohesion of a culture is based on the frequency of interactions with the group and the rarity of interactions with outsiders, not on any kind of higher-level influence on individuals.

Descriptions of customs, social norms, and cultures are always based on idealization and abstraction. Members of a community never have exactly the same ideas, preferences, or routines. That would be a miracle, given what is known about human learning and communication (Sperber 2006). There is always some variation among the members, no matter how comprehensive the socialization processes are. However, these idealized descriptions are still useful. They draw attention to features of the group that are typical and salient when it is contrasted with some other group.

Although communal properties, as I have described them, are tied to a social community defined by frequent interactions, the boundaries of these communities are fluid. This makes it possible to describe culture on various scales—for example, on the levels of a village, a local area, and a nation. However, descriptions on larger scales are bound to be more abstract and less rich in detail as individual variation takes its toll. The same flexibility that characterizes statistical and network properties applies also to communal properties, which can also be attributed to nonpersonal units. For example, it is possible to describe social norms that govern interactions between organizations.

When we consider communal properties as *idealizing abstractions from shared individual properties*, there is no need to refer to them as any kind of autonomous level of reality. They just describe more extensive facts than descriptions of the individual attitudes, habits, and preferences that constitute them. The scale perspective also appears natural when the explanatory use of communal properties is considered. For example, when we are explaining the behavior of an individual by appealing to social norms, we are referring to larger-scale facts about the group members that are causally relevant to the micro-level behavior. There is no need to postulate a separate realm of norms to understand what is happening. It is just that the expectations and responses of the other group members influence the individual's judgments about appropriate behavior.

2.3.4. Organizations and Their Properties

Organizations such as states, firms, parties, churches, and sport clubs are important parts of the social reality. While the community that is the basis for communal properties is not often clearly demarcated, a clear demarcation is often the case with organizations. They usually have specified criteria for membership, at least for the operational members. They also have rules that define the rights and duties of members and the roles of various functionaries. These (written or nonwritten) rules make it possible for organizations to have stability and continuity, so that it makes sense to talk about their continuing existence when their functionaries are replaced and the members change. Furthermore, many organizations exist (and are defined) in the context of other organizations, so one has to pay special attention to context when attempting to make sense of organizations.

Organizations as entities can have many properties that are not properties of their members. They can even have goals that are not the personal goals of their members, and some organizations are treated as legal persons. This has convinced many that organizations are real entities that should be treated as a separate ontological category. I do not have strong opinions about issues of ontological bookkeeping, as it is remembered that organizations are human artifacts that are always made of persons, their ideas about the rules, and often, of material artifacts. Whatever the organization does, is done by its members in its name. It is of crucial social importance whether an action, for example, a questionable comment, was made as a representative of an organization or as a private person. But these are facts about the status attributed to the behavior, not about the two completely different entities producing the behavior.

When a person causally interacts with an organization, she interacts with other persons (although this interaction is increasingly mediated via material artifacts such as ATM machines). There is no downward causal influence from a higher level. Everything happens at the same level; it is just that the intentional attitudes and relations of a larger group of people are important to the details of the local situation. Similarly, the influence of the organization on its members happens through other members, no matter how high up some of the members are in the organizational hierarchy. While the rules (and their interpretation by others) are external to any individual person, there is no need to posit them as a separate ontological category. These observations suggest that even in the case of organizations, the layer-cake model of the social world is not very illuminating. What is interesting about organizations is the habits and mental representations of their members, the resources they control as members of the organization, and their (materially mediated) interactions, not some higher ontological level.

Again it is good to return to real social scientific questions. They concern issues such as: How do large-scale collective enterprises—for example, organizations—manage (or fail) to achieve certain things? What kinds of unintended consequences do these collective activities have? How does a membership in such collective enterprises influence the individual members? The explanatory answers to these questions often refer to organizations and their properties, but there is no problem in conceiving them as large-scale things influencing smaller-scale things or other large-scale things.

These examples of macro social facts suggest a kind of flat view of society in which the difference between micro and macro is one of scale, not of different levels. The large-scale facts about distributions, frequencies, interactions, and relations have an irreducible explanatory contribution to make, but there is nothing comparable to the mind-brain relation. As a consequence, the metaphor of levels that underlies the layer-cake model does not really help to make sense of the issues that social scientists addressing social macro facts are facing. Giving it up will have a number of beneficial consequences.

First, there are some philosophical advantages. As I will argue in the next section, once we give up the image of levels, we get rid of the problem of causal exclusion that arises from the image of causally competing levels. There is no problem of downward causation as there are only causal influences from large-scale things to small-scale things and descriptions of large-scale things at various levels of abstraction. The problem is replaced with the more down to earth problem of explanatory selection: Under which description can we formulate the most robust claims about counterfactual dependence? Secondly, we no longer have to face the problem of finding an acceptable definition of the comprehensive individual level so that we can argue for or against methodological individualism. We can start analyzing real social scientific explanations instead and focus our attention on the possible contributions that large-scale things make to those on a smaller scale and what kinds of causal mechanisms mediate these influences.

This change in framing also has some advantages when considering relations between disciplines. The division of labor between psychology and the social sciences is justified by differences in scale and the importance of large-scale relations and interactions, not in terms of independent and autonomous levels of reality. This guarantees that the social sciences will never be reduced to psychological sciences. However, thinking in terms of scale also cuts down the false aspirations of disciplinary autonomy. When the social scientists are denied their own autonomous level of reality, the ideal of completely psychology-free social science becomes less appealing. It should be an empirical matter whether the details of human cognition matter for social explanation. It might be that is some cases it makes good mechanistic sense to incorporate some processes on the sub-personal level in the explanatory theory. I will return to this possibility in the final section.

2.4. Causation, Constitution, and Microfoundations

One prominent idea in the recent philosophy of biology debate about mechanisms has not been employed in the philosophy of social sciences debate.[1] This is the distinction between causation and constitution. Although the difference between constitutive and causal explanation has been noted earlier (Salmon 1984; see also Cummins 1983), it has only recently become a topic of systematic study (Craver 2007).

Both causation and constitution are relations of dependence (or determination), and they are easily confused. However, there are some crucial ontological differences. Causation is a relation between events; it is about changes in properties. Causation takes time, so we talk about causal processes. Finally, causation is characterized by the asymmetry of manipulation: The effect can be manipulated by manipulating the cause, but not the other way around (Woodward 2003).

In contrast, constitution relates properties. The properties (and relations) of parts constitute the properties of the system (sometimes also the relations to the environment are important). The whole is *made of* its parts and their relations. Unlike causation, constitution does not take time, and we do not talk about the process of constitution. Furthermore, the *relata* of constitution are not "independent existences" (as Hume called them). For this reason we cannot characterize the relation of constitution with the help of the asymmetry of manipulation. For example, the molecular structure of glass constitutes its fragility: To be fragile is to have a particular molecular structure; the fragility is not a consequence of the molecular structure. However, there is another sort of asymmetry: the asymmetry of existence. The parts preexist the system in the sense that the parts can exist independently of the system, but the system cannot exist independently of its parts (although the system can exist independently of particular parts).

An interesting sort of regress characterizes both causation and constitution. In the case of causation, we talk about chains of causation. This is based on the idea that for every event that is a cause, there is another event that is its cause. A similar idea applies to constitution; we assume that all parts can be further decomposed into their parts and their organization. We could call these chains of constitution. Now a tricky question is whether there exists a first cause that is not itself caused, and a similar problem can be stated concerning the ultimate building blocks of reality, but in this context we can leave them aside. There is no danger that such ultimate things will show up in the social sciences. However, these regress properties create chains of explanations, which are relevant from the point of view of the social sciences. The crucial thing in this context is to understand that although there is always an explanation for every social scientific explanatory factor, this does not imply that their explanatory status depends on us knowing the explanation for them. Both in the case of causation and constitution, an explanation presupposes that the *explanans* facts are the case, not that we have to have an explanation for those facts. I will return to this issue in the next section.

Explanation is about tracking relations of dependence. Although metaphysically the relations of constitution and causation are quite different, in terms of explanation the basic principles are quite similar. Both explanations attempt to track networks of counterfactual dependence. A causal explanation tells us how *the antecedent events* and *their organization* (timing and location) bring about the event to be explained. In contrast, a constitutive explanation describes how *the properties of the components* and *their organization* give rise to the system's properties.

In both cases we are looking for the difference-makers: The criterion of explanatory selection is counterfactual. As the precise *explanandum* is best characterized in contrastive terms (why x is the case rather than x*), we are interested in the differences that would have made the difference we are interested in (Woodward 2003; Ylikoski 2007; Northcott this volume). In the case of causation these differences are in antecedent events; in the case of constitution these differences are in the properties of parts (or in their organization). Also in both cases it makes sense to ask a further question: Why does the counterfactual dependence hold? The answers to these questions will in both cases draw from the same body of mechanical knowledge, so it is understandable that in the philosophy of biology debates both explanations are called mechanical explanations. So, despite the important metaphysical differences, the same basic ideas about explanation can be applied to both cases.

Not only are the principles of explanatory relevance similar, so are the explanatory questions. This leads easily to confusion. Consider the question: "Why is this glass fragile?" The question is ambivalent: It could either mean "How did the glass become fragile?" or it could mean "What makes the glass fragile?" The first question is causal; the latter question constitutive. The answer to the causal question will tell us about the causal history of the glass—it will specify the crucial features of the process that led to the object being fragile rather than robust. The answer to the constitutive question will not focus on earlier events. It will detail the relevant aspects of the object's molecular structure that makes it fragile. So while the explanation-seeking questions may look the same, the request for explanatory information is quite different. Without a clear understanding of the differences between causation and constitution, some confusion is bound to occur. This is also the case in philosophy of social sciences. For example, it is quite a different thing to explain how a regime became stable than to explain what makes it stable. While some of the facts cited by both explanations might be the same, they are addressing different *explananda*: One is focused on how the causal capacity was acquired and the other on the basis of that causal capacity. A social scientist is usually interested in both questions, but she should not confuse them with each other.

For all social macro properties, one can ask both constitutive and causal why- and how-questions. (Although for some statistical properties the constitutive questions are relatively trivial.) The first sort of questions asks how the macro properties are constituted by the micro-level entities, activities, and relations. The aim is to track how the details of macro-level facts depend on the micro details. The question is often how the macro facts would have been different if some of the micro facts had been different in some specific way. These questions can also be characterized in terms of interventions: How would the macro facts change if some of the micro facts were changed? Notice that here intervention is a causal notion (all change happens in time), but the dependence of interest is constitutive.

A clear example of constitutive explanation is an explanation for the difference in the problem-solving capacities of two groups. The crucial difference might be in the properties of the members, such as their intelligence or social skills. Alternatively, the pivotal factors might be the informal social norms that characterize the interactions within the group or its formal organization. Of course, the explanation may also be found in some combination of these factors. Just like in this example, the usual *explananda* of constitutive explanations are causal capacities and dispositions of the whole. The constitutive explanation tells us what gives the whole (population, group, organization, or society) those properties, and the answer is found in the causal capacities of the parts and their organization.

The *explanantia* in constitutive explanations are always at the micro level. As the explanation attempts to capture what the whole is made of, an appeal to the properties of the whole does not really make sense. In this sense, the methodological individualists, and other reductionists, have been on the right track. On the other hand, the explanation of macro properties does not in any way diminish their reality: The wholes are as real as their parts. This implies that those methodological individualists who have suggested that a micro explanation somehow eliminates the macro properties are either metaphysically confused or just choosing their words badly. The talk about macro reducing to micro makes as little sense as the talk about reducing effects to their causes.

The causal questions about the macro social properties are concerned with their origin, persistence, and change. These explanations are tracking counterfactual dependencies between events. How would have the outcome been different if some of the causes had been different in some specified manner? What kind of difference would an intervention on some antecedent facts make? The *explanantia* in these causal explanations are always antecedent events.

This is the context in which confusion between constitution and causation can create trouble. If we are considering simple causal statements about causal dependence, individualists tend to make the claim that the causes have to be at the micro level. However, nothing in the notion of causation implies that the real causal work is always to be found at the micro level. Of course, the notion of constitution implies that every time we have a cause at a macro level, we also have micro level facts that constitute it. If we stick to the counterfactual criterion of explanatory selection, as I think we should, there is no a priori reason to privilege micro-level causes (Woodward 2003, 2008). It is sufficient that there is an appropriate counterfactual dependence between the macro variable and the *explanandum*. Of course, in many cases the justification of a claim about this causal dependence might require some knowledge of the underlying mechanisms. However, this observation about the justification of a causal claim should not be confused with the claim itself. Similarly, although adding mechanistic details to the explanation will involve references to micro-level processes, this does not imply that the macro facts will lose their explanatory relevance. They will still be possible difference-makers and legitimate explanatory factors. In other words, although the information about the relevant mechanistic details improves the explanation significantly, it does not remove the causal relevance of the initial invariance involving macro-level facts.

In the counterfactual account of causal relevance, the location of explanatory relevance at the micro or macro level is a contingent matter that depends on the *explananda* that one is addressing. There is no reason to assume that the most invariant counterfactual dependence (with respect to the contrastively specified *explanandum*) will always be found at the micro level. Similarly, one has to give up the often presented suggestion that levels of explanation should match so that macro would always explain macro and micro would always explain micro. The issues of explanatory relevance (how the explanatory factors are selected, at which level of abstraction they are described, etc.) are always determined by the facts of the case and the details of the intended *explanandum*, not by generic philosophical arguments.

2.4.1. The Proper Role of Microfoundations

Is the above argument about the legitimacy of macro-level causal facts compatible with the mechanistic call for microfoundations? I want to argue that it is fully compatible with the core ideas of mechanism-based thinking. Contrary to the common assumption, the point of mechanistic microfoundations is not that we have more real causes at the micro level, but to have a better grasp of the explanatory dependence underlying the causal relation involving macro variables. Consequently, the advocates of mechanism-based explanations should not call into question the reality of macro-level causal relations. Instead, they should emphasize the importance of microfoundations for understanding these dependencies. There are a number of reasons why microfoundations are important.

First, all causal relations involving macro properties are mechanism-mediated causal relations. Understanding how the dependence involving macro variables is constituted helps to understand why that particular dependence holds (Ylikoski 2011). It also integrates the piece of causal information contained in the macro-level generalization to other pieces of explanatory knowledge (Ylikoski and Kuorikoski 2010). This is certainly a form of explanatory understanding that we should be interested in if we take the notion of explanatory social science seriously.

However, the utility of this information is not limited to the expanded theoretical understanding. It also often tells about the conditions under which the causal dependence in question will hold. There are three dimensions to this knowledge. First, there is knowledge about the range of values of the *explanandum* variable that are possible without the dependence breaking apart. Second, there is knowledge about the sensitivity of the dependence to changes in background conditions. Finally, there is possible knowledge about alternative interventions that could bring about similar effects. Without knowledge of these issues, the explanatory use of the macro-level explanatory generalization can be very risky business. It is very difficult to extrapolate to other cases without understanding the background mechanisms (Ylikoski 2011; see also Cartwright, this volume, Kincaid, this volume).

Apart from an expanded understanding and the security of an explanatory claim, the insight into the underlying mechanisms might also help to improve the explanatory generalization. With the help of a mechanistic understanding, one might be able to make the *explanandum* more precise or to reformulate the explanatory generalization in such a manner that it allows a broader range of values of the *explanandum* variables or background conditions (Ylikoski 2011).

These considerations justify the presumption that microfoundations are important for proper explanatory understanding. However, they do not demolish the explanatory relevance of macro facts. On the contrary, they put them in the right context as the mechanisms bridge the large-scale micro facts to causal interactions between persons and to their decision-making processes. I think this is the point James Coleman (1990) attempted to make with his often misunderstood graph.

Following Hedström and Swedberg (1998, 23), I refer to the arrows in figure 2.1 as situational mechanisms (arrow 1), action-formation mechanisms (arrow 2), and

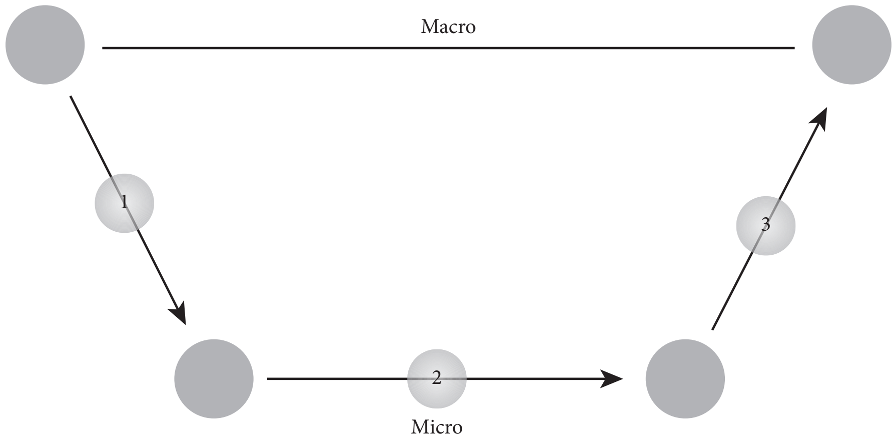


Figure 2.1 Macro-Micro Linkages

transformational mechanisms (arrow 3). The situational mechanisms describe how social structures constrain individuals' actions and cultural environments shape their desires and beliefs, the action-formation mechanisms describe how individuals choose their preferred courses of action among the feasible alternatives, and the transformational mechanisms describe how individual actions produce various intended and unintended social outcomes.

Coleman was critical of nonmechanistic explanations that remain at the level of macro regularities. However, there is no reason to assume that he was denying the causal relevance of macro social facts. Rather, his point was to make it clear that proper sociological understanding requires that we understand both the mechanisms by which large-scale social facts influence the local decision-making processes of individual agents (the situational mechanisms) and the mechanisms by which individual actions create and influence macro social facts (the transformational mechanisms). He was calling for mechanisms that bridge the levels, not just descriptions that somehow reduce the macro facts to individual level facts. Only when we understand the relevant mechanisms, do we have a satisfactory theoretical grasp of the social phenomena in question.

Coleman's criticism of Weber's (partial) explanation of the emergence of modern capitalism in Western Europe illuminates these points. Weber started with an idea that was commonplace in late nineteenth-century Germany: There is a close connection between Protestantism, entrepreneurism, and the rise of capitalism. To substantiate this vague explanatory suggestion, Weber asked what changes the emergence of Protestantism brought about in the beliefs, desires, and communal practices of individual agents. This question has both causal and constitutive dimensions that are not clear in Coleman's analysis. However, Coleman's focus is on Weber's second causal question: How did these changed life practices of individuals influence economic activities and institutions and how did these changes in turn facilitate the formation of modern capitalism? Coleman's central point was that Weber was not clear enough about this last passage of the causal chain. He was not able to give a sufficiently clear account of the transformative mechanisms that connected the Protestant ethic to the rise of modern capitalism. In other words, Weber was not able to show how the changes at the micro level (the life practices of Protestants) bought about a major macro-level outcome (the early forms of modern capitalism). As the crucial mechanism is lacking, so is the legitimacy of Weber's causal claim about history.

Here it is important to see the difference between the justificatory and explanatory roles of mechanisms. Coleman's analysis shows why it is legitimate to challenge Weber's causal claim. Knowledge of the causal mechanisms have an important role in the justification of historical causal claims, so pointing to the missing details of the causal chain constitutes a challenge to the legitimacy of the causal claim. However, this criticism of a singular causal claim does not imply that Coleman generally considers macro-level facts to be nonexplanatory or causally impotent. He is simply challenging the justification of this particular historical hypothesis.

2.5. Intentional Fundamentalism

Arguments for methodological individualism often appeal to the special explanatory status of intentional explanations. I call this position intentional fundamentalism. According to intentional fundamentalism, the proper level of explanation in the social sciences is the level of the intentional action of individual agents. The intentional fundamentalist assumes that explanations given at the level of individual action are especially satisfactory, fundamental, or even ultimate. In contrast to explanations that refer to supra-individual social structures, properties, or mechanisms, there is no need to provide microfoundations for intentional explanations. They provide rock-bottom explanations. In other words, according to intentional fundamentalism, the intentional explanations of individual actions are *privileged* explanations.

Although intentional fundamentalism can take various forms, it is often related to rational choice theory. French social theorist Raymond Boudon (1998, 177) expresses the idea clearly: "When a sociological phenomenon is made the outcome of individual reasons, one does not need to ask further questions." The idea is that in the case of supra-individual explanations there is always a black box that has to be opened before the explanation is acceptable, but in the case of intentional explanation there is no such a problem: "The explanation is final" (Boudon 1998, 172). Diego Gambetta appeals to same sort of finality (1998, 104): "Not only will a rational choice explanation be parsimonious and generalizable; it will also be the end of the story."[2]

My claim in this section is that intentional fundamentalism is not compatible with the causal mechanistic account of explanation. As intentional fundamentalism is often advocated by rational choice theorists and as many believe that rational choice explanations are the best examples of mechanical explanations in the social sciences, this incompatibility claim is of some interest. If my argument is valid, it suggests that the relation of rational choice theory and a mechanism-based philosophy of science requires some rethinking. It also implies that one common argument for methodological individualism is much less credible than is commonly assumed.

2.5.1. The Regress Argument

To make sense of intentional fundamentalism, we should start with *the explanatory regress argument for methodological individualism*. Methodological individualists often make the case that nonindividualist explanations are either explanatorily deficient or not explanatory at all. At most, they allow that explanations referring to macro social facts are placeholders for proper (individualistic) explanatory factors. In this view, the explanatory contribution of supra-individual explanations is at best derived: They are explanatory because they are (in principle) backed up by a truly explanatory story. This is the regress of explanations argument: Unless grounded at the lower level, explanations at the macro level are not acceptable. The underlying general principle is the following:

> [P] A genuine explanation requires that the *explanans* is itself explained or is self-explanatory.

In short, the explanatory buck has to stop somewhere.

The principle [P] is general, and it raises the possibility of an explanatory regress that is only halted at a fundamental (physical) level. This would be highly unintuitive, so for the intentional fundamentalist the buck stops at the level of (self-interested) rational intentional action. This level is treated as inherently understandable, as shown in the above quotations from Boudon. The inherent intelligibility of intentional action explains why the search for microfoundations should stop at the level of the individual. The special status of intentional explanation also makes the explanatory regress argument safe for the methodological individualist: He can use the argument's full force against anti-individualists who cannot make a similar claim about a privileged status, and it does not challenge the legitimacy of his favored explanatory factors.

The fundamentalist argument for individualism fails for a number of reasons. The first reason is that the principle [P] is not valid. The explanatory relation between the *explanans* and the *explanandum* is independent from the question of whether the *explanans* is itself explained. An explanation of X in terms of Y presupposes that Y is the case, but it does not presuppose that Y is itself explained. Of course, it would be great also to have an explanation for Y, but this is a separate issue from the legitimacy of the explanatory relationship between Y and X. The distinctness of these issues implies that the regress does not begin.

Why whould anyone believe in [P]? One plausible suggestion is the following: The belief in [P] arises from a straightforward confusion between justification-seeking and explanation-seeking why-questions. It makes sense to ask how well justified are those things that one appeals to in justification of one's beliefs. It also makes sense to ask whether one is justified in believing the things that one appeals to in one's explanation. However, justifying one's belief in Y is not the same as explaining why Y is the case.

2.5.2. Intentional Explanations without a Special Status

Another reason for the failure of the regress argument is that intentional explanations lack the special properties assumed by the argument. If one accepts the mechanistic account of explanation, as many advocates of rational choice sociology do, such a special status does not make any sense. The assumption that human deliberation is a black box that should not be opened is more in line with nineteenth-century hermeneutic romanticism than with causally oriented social science. Of course, the chain of mechanistic explanations will end somewhere (if there is such a thing as a fundamental level), but that stopping point is not the level of individual rational action.

A mechanistic explanation appeals to micro-level processes, but nothing in the notion of mechanistic explanation implies that these micro things would always be facts about the intentional actions of individuals. Mechanisms that cite supra-individual entities or properties are certainly possible (Mayntz 2004). For example, various filtering mechanisms that are analogical to natural selection are difficult to understand other than as population-wide processes, and when the units that are selected are organizations (for example, firms), it is natural to conceive the mechanism as supra-individual. Similarly, the crucial parts of the explanatory mechanism could well be located below the level of intentional psychology. For example, various facts about human information processing—for example, implicit biases (see Kelly and Mallon, this volume)—could well be relevant for explanatory understanding of intentional action. There is no valid reason to give up mechanistic thinking in the case of intentional action.

Another reason to challenge intentional fundamentalism is the implicit realism of mechanistic thinking. For mechanists, explanation is factive. It is not enough that the explanation saves the phenomenon: It should also represent the essential features of the actual causal structure that produces the observed phenomena. So, if the explanation refers to the goals, preferences, or beliefs of agents, the agents should indeed have those mental states. Mere as-if storytelling does not suffice for a mechanistic explanation as it does not capture the relevant parts of the causal process. This realist attitude goes against the instrumentalist attitude common among many rational choice theorists. The fact that one can rationalize any behavior does not imply that those rationalizations are also the correct causal explanations for those behaviors. Similarly, the human fluency in coming up with intentional accounts for our behavior is not a reason for regarding them as superior explanations.

It is important to understand the limited nature of my argument. I am not denying that intentional explanations are, and will be, an indispensable part of the social scientific explanatory repertoire. For me, intentional explanations are legitimate causal explanations. Furthermore, the intentional attitudes of individuals play an important role in most mechanism-based explanations of social phenomena. The only thing I am challenging is the supposed special explanatory status of intentional or rational accounts of human action. In the mechanistic account of explanation, the importance of certain sorts of explanatory factors is not a basis for their privileged status.

Neither should my rejection of intentional fundamentalism be regarded as a wholesale attack on the use of rational choice theory. For many social scientific purposes, a rather simple version of intentional psychology is both preferable and sufficient. For example, when one is attempting to make sense of social complexity, it is understandable that social scientists attempt to keep the psychological assumptions of their models very simple. Such idealizations are fully legitimate if they do not lead to a gross misrepresentation of the causal mechanism under consideration. However, the practical necessity of these idealizations does not constitute a justification for accepting intentional fundamentalism.

Furthermore, my argument should not be regarded as an argument against the claim that there should exist a division of labor between the social sciences and the sciences of cognition. However, it follows from the flexibility of mechanistic levels that the boundaries of this division of labor are adjustable and not fixed. It is inherent in the idea of mechanistic explanation that all the gaps between levels of analysis are ultimately to be bridged by mechanistic interfield theories. So the challenge for the social sciences is not to define their objects of study in such a way that they are in no way touched by psychological sciences, but to look at ways in which social and cognitive mechanisms can be meaningfully combined. This is not as easy as it sounds, as recent attempts to combine neuroscience and economics show (Kuorikoski and Ylikoski 2010).

2.6. Conclusions

In this chapter, I have attempted to show what consequences the mechanism-based account of explanation would have on issues traditionally discussed under the title of methodological individualism. Borrowing some ideas developed by philosophers who have studied the mechanistic explanation in the biological sciences, I have argued that we should give up the notion of a unique, privileged, and comprehensive individual level that has been a presupposition of the individualism debates. In addition, I have argued that rather than employing metaphors borrowed from the philosophy of mind for micro-macro relations, we should pay closer attention to how real macro social facts figure in social scientific theories and explanations. There the micro-macro issue is more an issue of bridging large-scale social facts to small-scale social interactions rather than that of finding a way to see relations between autonomous levels of reality.

NOTES

1. There are some exceptions. For example, Wendt (1998) distinguishes between causation and constitution. However, his discussion of constitution is very confused. His notion of constitution covers not only the constitution of causal capacities, but also causal preconditions, definitions, and other conceptual relations. The standard philosophy of science notion that I am using is limited only to the constitution of causal capacities.

2. The key issue here is not whether these authors would ultimately subscribe to intentional fundamentalism. I am only claiming that in these passages they argue as if intentional fundamentalism is correct.

REFERENCES

Bechtel, William. 2006. *Discovering Cell Mechanisms. The Creation of Modern Cell Biology*. New York: Cambridge University Press.

Bechtel, William. 2008. *Mental Mechanism. Philosophical Perspectives on Cognitive Neuroscience*. London: Routledge.

Boudon, Raymond. 1998. "Social Mechanisms Without Black Boxes." In *Social Mechanisms: An Analytical Approach to Social Theory*, Peter Hedström and Richard Swedberg, eds., 172–203. Cambridge: Cambridge University Press.

Burt, Ronald. 1992. *Structural Holes: The Social Structure of Competition*. Cambridge, MA: Harvard University Press.

Coleman, James. 1990. *Foundations of Social Theory*. Cambridge, MA: The Belknap Press.

Craver, Carl. 2007. *Explaining the Brain: Mechanisms and the Mosaic Unity of Neuroscience*. Oxford: Clarendon Press.

Cummins, Robert. 1983. *The Nature of Psychological Explanation*. Cambridge, MA: Bradford/The MIT Press.

Darden, Lindley. 2006. *Reasoning in Biological Discoveries. Essays on Mechanisms, Interfifield Relations, and Anomaly Resolution*. Cambridge: Cambridge University Press.

Elster, Jon. 1989. *Nuts and Bolts for the Social Sciences*. Cambridge: Cambridge University Press.

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

Gambetta, Diego. 1998. "Concatenations of Mechanisms." In *Social Mechanisms: An Analytical Approach to Social Theory*, Peter Hedström and Richard Swedberg, eds., 102–24. Cambridge: Cambridge University Press.

Harré, Rom. 1970. *The Principles of Scientific Thinking*. London: Macmillan.

Hedström, Peter. 2005. *Dissecting the Social. On the Principles of Analytical Sociology*. Cambridge: Cambridge University Press.

Hedström, Peter, and Peter Bearman, eds. 2009. *The Oxford Handbook of Analytical Sociology*. Oxford: Oxford University Press.

Hedström, Peter, and Richard Swedberg. 1998. "Social Mechanisms: An Introductory Essay." In *Social Mechanisms: An Analytical Approach to Social Theory*, Peter Hedström and Richard Swedberg, eds., 1–31. Cambridge: Cambridge University Press.

Hedström, Peter, and Petri Ylikoski. 2010. "Causal Mechanisms in the Social Sciences." *Annual Review of Sociology* 36: 49–67.

Hempel, Carl. 1965. *Aspects of Scientific Explanation*. New York: The Free Press.
Horgan, Terence. 1993. "From Supervenience to Superdupervenience: Meeting the Demands of a Material World." *Mind* 102: 555–86.
Kim, Jaegwon. 1993. *Supervenience and Mind*. Cambridge: Cambridge University Press.
Kim, Jaegwon. 1998. *Mind in a Physical World*. Cambridge, MA: The MIT Press.
Kincaid, Harold. 1996. *Philosophical Foundations of the Social Sciences. Analyzing Controversies in Social Research*. Cambridge: Cambridge University Press.
Kuorikoski, Jaakko. 2009. "Two Concepts of Mechanism: Componential Causal System and Abstract Form of Interaction." *International Studies in the Philosophy of Science* 23 (2): 143–60.
Kuorikoski, Jaakko, and Petri Ylikoski. 2010. "Explanatory Relevance Across Disciplinary Boundaries—The Case of Neuroeconomics." *Journal of Economic Methodology* 17 (2): 219–28.
Little, Daniel. 1991. *Varieties of Social Explanation: An Introduction to the Philosophy of Social Science*. Boulder, CO: Westview Press.
Mayntz, Renate. 2004. "Mechanisms in the Analysis of Social Macro-Phenomena." *Philosophy of the Social Sciences* 34 (2): 237–59.
McCauley, Robert. 2007. "Reduction: Models of Cross-Scientific Relations and Their Implications for the Psychology-Neuroscience Interface." In *The Handbook of Philosophy of Science. Philosophy of Psychology and Cognitive Science*, Paul Thagard, ed., 105–58. Amsterdam: North Holland.
Oppenheim, Paul, and Hilary Putnam. 1958. "Unity of Science as a Working Hypothesis." In *Concepts, Theories, and the Mind—Body Problem, Minnesota Studies in the Philosophy of Science II*, Herbert Feigl, Michael Scriven, and Grover Maxwell, eds., 3–36. Minneapolis: University of Minnesota Press.
Richardson, Robert. 2007. "Reduction without the Structures." In *The Matter of the Mind. Philosophical Essays on Psychology, Neuroscience, and Reduction*, Maurice Schouten and Huib Looren de Jong, eds., 123–45. Oxford: Blackwell.
Salmon, Wesley. 1984. *Scientific Explanation and the Causal Structure of the World*. Princeton, NJ: Princeton University Press.
Sawyer, R. Keith. 2005. *Social Emergence. Societies as Complex Systems*. Cambridge: Cambridge University Press.
Scott, John. 2000. *Social Network Analysis. A Handbook*. 2nd ed. London: Sage.
Sperber, Dan. 2006. *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell.
Thagard Paul. 1999. *How Scientists Explain Disease*. Princeton, NJ: Princeton University Press.
Turner, Stephen. 1994. *The Social Theory of Practices. Tradition, Tacit Knowledge and Presupposition*. Cambridge: Polity Press.
Udéhn, Lars. 2001. *Methodological Individualism: Background, History and Meaning*. London: Routledge.
Wendt, Alexander. 1998. "On Constitution and Causation in International Relations." *Review of International Studies* 24 (5): 101–17.
Wimsatt, William. 2007. *Re-Engineering Philosophy for Limited Beings. Piecewise Approximations to Reality*. Cambridge, MA: Harvard University Press.
Woodward, James. 2002. "What Is a Mechanism? A Counterfactual Account." *Philosophy of Science* 69 (3): S366–S377.
Woodward, James. 2003. *Making Things Happen. A Theory of Causal Explanation*. Oxford: Oxford University Press.

Woodward, James. 2008. "Mental Causation and Neural Mechanisms. In Being Reduced." *Essays on Reduction, Explanation, and Causation*, eds. J. Hohwy and J. Kallestrup, 218–62. Oxford: Oxford University Press.
Wright, Cory, and William Bechtel. 2007. "Mechanisms and Psychological Explanation." In *The Handbook of Philosophy of Science. Philosophy of Psychology and Cognitive Science*, ed. Paul Thagard, 31–79. Amsterdam: North Holland.
Ylikoski, Petri. 2003. "Explaining Practices." *Protosociology* 18: 316–30.
Ylikoski, Petri. 2007. "The Idea of Contrastive Explanandum." In *Rethinking Explanation*, eds. Johannes Persson and Petri Ylikoski, 27–42. Dordrecht: Springer.
Ylikoski, Petri. 2011. "Social Mechanisms and Explanatory Relevance." In *From Social Mechanisms to Analytical Sociology*, ed. Pierre Demeulenaere, 154–72. Cambridge: Cambridge University Press.
Ylikoski, Petri, and Jaakko Kuorikoski. 2010. "Dissecting Explanatory Power." *Philosophical Studies* 148 (2): 201–19.
Zahle, Julie. 2006. "Holism and Supervenience." In *The Handbook of Philosophy of Science. Philosophy of Anthropology and Sociology*, eds. Stephen P. Turner and Mark Risjord, 311–41. Amsterdam: North Holland.

CHAPTER 3

MECHANISMS, CAUSAL MODELING, AND THE LIMITATIONS OF TRADITIONAL MULTIPLE REGRESSION

HAROLD KINCAID

My target in this chapter are three things: the idea that the social sciences need mechanisms, a standard way multiple regression is used in the social sciences to infer causality, and the usefulness of the directed acyclic graph (DAG) approaches (Sprites, Glymour, and Scheines, 2001; Pearl 2000) in understanding social causality. Philosophers of science as well as social scientists have often claimed that the social, behavioral, and biomedical sciences need mechanisms (Elster 1983; Hedström and Swedberg 1998). However, this claim is often muddled. The idea of a mechanism is often unexplained or used in different senses. The reason mechanisms given for why we need mechanism are various, often left inexplicit, and not related to the sense of mechanism at play. In this chapter I use work on causal modeling with directed acyclic graphs to show some circumstances where mechanisms are needed and not needed and to give clear reasons why that is the case. In the process I show how standard regression practices in the social sciences can go wrong and how they can be improved. I also point to some limitations of the DAG program in identifying mechanisms in the social sciences. My examples come from development economics.

3.1. A Standard Practice and a Common Demand

Here is a standard practice in social research:[1] An investigator is interested in some outcome variable *Y*. Data sets are collected or existing data sets identified which have measurements of *Y* and other variables $X_1 \ldots X_n$ that might be associated with *Y*. The $X_1 \ldots X_n$ variables are called the independent variables and described as the determinants of *Y* or as the factors associated with *Y*. Actual claims to causation are studiously avoided as results are reported. Multiple regressions are run on the data set, producing estimates of the coefficient sizes on the independent variables and providing statistical significance levels of those variables. The significance tests are then used as indicators of whether each independent variable is truly a determinant or factor in outcome *Y*. Sometimes variables are kept in the regression only if they are significant and the coefficients on the variables that remain are then re-estimated. The regression coefficients are taken to be a measure of the size of the factor or determinant. Proper caution is exercised in reporting the results by noting that correlation is not causation. However, in the closing section where the importance of the results are discussed, policy implications are noted, throwing caution to the wind for causal conclusions about what would happen if we could intervene.

This practice is widespread across the social sciences. A particularly vibrant example that I will return to later in the chapter is work in economics and development studies on growth using cross-country regressions. I have in mind, for example, Robert Barro's *The Determinants of Economic Growth* (1998). Data are collected on most countries in the world. Each country is treated as an individual data point. GDP per capita or some related variable is treated as the outcome variable. The independent variables are observations on each country concerning economic and noneconomic determinants. A regression equation of the following form is estimated:

$$\text{gdp} = \text{investment} + \text{open markets} + 57 \text{ other variables}$$

The 57 other variables do not refer to Heinz, the American brand of ketchup. Rather, the total number of initial independent variables in this country regression work is generally actually 59. Some of those 59 are motivated by economic theory as investment level obviously is. However, probably the majority are variables that someone thought might somehow be relevant and variables for which we have data. So, for example, religion is always included.

Significance levels are reported, variables dropped, and regressions rerun, producing papers with titles such as "I Just Ran 2 Million Regressions" (Sala-i-Martin 1997).[2] Different studies end up with different sets of variables in the final regression. In Barro's work the final set includes common economic variables and institutional variables that fit with the Washington Consensus of the 1990s that emphasized open markets, minimal sized states, and protection of property rights among other things. Interestingly, the regressions run by Sala-i-Martin and by Hoover and Perez (2004) find the education variable to be nonsignificant. Policy recommendations are drawn based on the final surviving regression.

Cross-country growth regressions are not an outlier in social science research. The same kinds of practices are repeated again and again in the social sciences as well as in such fields as epidemiology and public health. For example, a major concern across economics and sociology is the determinants of inequality and the distribution of wealth and income across individuals. A great many studies have been published reporting regressions results using the same recipe as found in the cross-country regressions (Bowles, Gintis, and Groves 2005 is a typical example). Standard analytic epidemiological studies of disease outcomes paired with a final set of covariates or risk factors and their associated coefficients do something similar, though epidemiologists are more wary of extensive stepwise regression (Kincaid 2011a).

These standard uses of regressions to infer causes is one my concerns in this chapter.[3] Another is the demand made by both social scientists and philosophers that good social research must produce mechanisms. This idea predates the current interest in mechanisms by philosophers of science (Machamer, Darden, and Craver 2000). Elster, in *Explaining Technological Change* (1983), argued that we need mechanisms in terms of individual behavior to identify spurious correlations in the social sciences. All the authors in Hedström and Swedberg's *Social Mechanisms* argue that mechanisms should be central to social theory.

I am suspicious of any blanket claim about mechanisms in the social sciences for two reasons. First, I am suspicious of broad methodological pronouncements in science in general. In practice methodological rules require domain and context specific knowledge for their interpretation and application (Day and Kincaid 1994; Kincaid 2011b). Simplicity, for example, has a role in science, but the work that it does often comes from domain-specific instantiations that embody substantial empirical claims (Sober 1989). I would expect the same for claims about mechanisms in the social sciences. A second reason for skepticism about the demand for mechanism results from the fact that the claim can be given several different readings and motivations that are logically independent and need not stand or fall together. A framework for thinking about those differences would help clarify the issues, and I turn now to sketch out the logical space of claims.

A first question is what we want mechanisms for. As I argued some time ago (1996, 1997), we might want a mechanism for *explanatory* purposes or for providing *evidence*. These need not be the same. I may have a well-confirmed association or even a causal claim, but think it is not sufficiently deep enough to explain in some sense—that to explain I need to know how the relation obtains. On the other hand, I might want mechanisms because I doubt an association is real in the first place and believe that providing a mechanism would lend it further credibility. So explaining and confirming with mechanisms can come apart.

Though I will not emphasize it much below, the confirmation versus explanation dichotomy does not exhaust the things we might want to do with mechanisms.

If we move to a more dynamic situation where the generation of research questions and hypotheses is our interest, then mechanisms might play a role there that goes beyond confirmation or explanation. Clearly these sorts of uses are at work in the descriptive accounts of mechanisms in scientific practice from philosophers of science (Bechtel and Richardson 2010).

We can also use mechanisms to confirm two different types of causal claims: assertions that a causal relation *exists* versus assertions about the *size* of the relationship. It is one thing to know that C causes E, another to know how changes in the various values of C result in differing values of E. Mechanisms might be valuable for determining effect size but not effect or vice versa. So having a mechanism might increase my evidence that changes in the interest rate cause changes in employment or it may be needed for me to infer how much a change in interest rates increases or decreases employment.

In terms of explanation or understanding, we can likewise have distinct goals. Mechanisms might help us with the purely social scientific goal to have a *theoretical* understanding of the social phenomena. Achieving that goal does not necessarily mean we know how to *intervene* successfully to change outcomes; mechanisms might be more important in the latter case than in the former. A randomized clinical trial might show us that treatment C causes E without knowing the mechanism. But if we want to intervene, we might want to know the process whereby C causes E, for it is possible our intervention might bring about C and at the same time block the process producing the effect.

Another important difference in thinking about mechanism turns on whether we want *horizontal* or *vertical* mechanisms (see figure 3.1). Asking for horizontal mechanisms is asking for the steps that led from C to E—the intervening causes that makes for a continuous process. I label these horizontal because they are mechanisms at the same level as what the mechanisms relate. The case represented in the figure is the simplest case. In more complex cases M has itself other causal relations at the same level. I will call these more complex causal relations *causal structures*. Horizontal mechanisms are then either simply intervening variables or intervening variables plus causes they interact with over and above C and E.

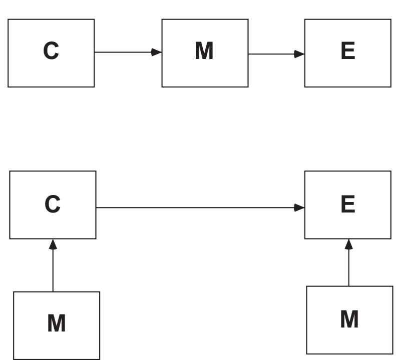


Figure 3.1 Horizontal versus vertical mechanisms