Communicative Action and Rational Choice

Joseph Heath

Communicative Action and Rational Choice

Joseph Heath

The MIT Press
Cambridge, Massachusetts
London, England

Library of Congress Cataloging-in-Publication Data

Heath, Joseph, 1967–
Communicative action and rational choice / Joseph Heath.
p. cm.—(Studies in contemporary German social thought)
Includes bibliographical references and index.
ISBN 0-262-08291-8 (hardcover : alk. paper)
1. Habermas, Jürgen. 2. Ethics. 3. Act (Philosophy) 4. Communication—Philosophy. I. Title. II. Series.
B3258.H324 .H43 2001
170—dc21 00-032918

Communicative Action and Rational Choice

Studies in Contemporary German Social Thought (partial listing)
Thomas McCarthy, general editor

Seyla Benhabib and Fred Dallmayr, editors, *The Communicative Ethics Controversy*
James Bohman, *Public Deliberation: Pluralism, Complexity, and Democracy*
James Bohman and Matthias Lutz-Bachmann, editors, *Perpetual Peace: Essays on Kant's Cosmopolitan Ideal*
Craig Calhoun, editor, *Habermas and the Public Sphere*
Jean Cohen and Andrew Arato, *Civil Society and Political Theory*
Maeve Cooke, *Language and Reason: A Study of Habermas's Pragmatics*
Helmut Dubiel, *Theory and Politics: Studies in the Development of Critical Theory*
John Forester, editor, *Critical Theory and Public Life*
Jürgen Habermas, *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*
Jürgen Habermas, *Justification and Application: Remarks on Discourse Ethics*
Jürgen Habermas, *On the Logic of the Social Sciences*
Jürgen Habermas, *The Inclusion of the Other: Studies in Political Theory*
Jürgen Habermas, *Moral Consciousness and Communicative Action*
Jürgen Habermas, *The New Conservatism: Cultural Criticism and the Historians' Debate*
Jürgen Habermas, *The Philosophical Discourse of Modernity: Twelve Lectures*
Jürgen Habermas, *Philosophical-Political Profiles*
Jürgen Habermas, *Postmetaphysical Thinking: Philosophical Essays*
Jürgen Habermas, *On the Pragmatics of Communication*
Jürgen Habermas, *On the Pragmatics of Social Interaction: Preliminary Studies in the Theory of Communicative Action*
Jürgen Habermas, *The Structural Transformation of the Public Sphere: An Inquiry into a Category of Bourgeois Society*
Jürgen Habermas, editor, *Observations on "The Spiritual Situation of the Age"*
Joseph Heath, *Communicative Action and Rational Choice*
Axel Honneth, *The Critique of Power: Reflective Stages in a Critical Social Theory*
Axel Honneth, *The Struggle for Recognition: The Moral Grammar of Social Conflicts*
Axel Honneth and Hans Joas, editors, *Communicative Action: Essays on Jürgen Habermas's* The Theory of Communicative Action
Axel Honneth, Thomas McCarthy, Claus Offe, and Albrecht Wellmer, editors, *Cultural-Political Interventions in the Unfinished Project of Enlightenment*
Axel Honneth, Thomas McCarthy, Claus Offe, and Albrecht Wellmer, editors, *Philosophical Interventions in the Unfinished Project of Enlightenment*
Elliot L. Jurist, *Beyond Hegel and Nietzsche: Philosophy, Culture, and Agency*
Michael Kelly, editor, *Critique and Power: Recasting the Foucault/Habermas Debate*
Cristina Lafont, *The Linguistic Turn in Hermenutic Philosophy*
Maurizio Passerin d'Entrèves and Seyla Benhabib, editors, *Habermas and the Unfinished Project of Modernity: Critical Essays on* The Philosophical Discourse of Modernity
Kirk Pillow, *Sublime Understanding: Aesthetic Reflection in Kant and Hegel*
William E. Scheuerman, *Between the Norm and the Exception: The Frankfurt School and the Rule of Law*
Albrecht Wellmer, *Endgames: Essays and Lectures on the Irreconcilable Nature of Modernity*
Albrecht Wellmer, *The Persistence of Modernity: Essays on Aesthetics, Ethics and Postmodernism*
Rolf Wiggershaus, *The Frankfurt School: Its History, Theories, and Political Significance*

For my mother and father

Preface

This book has been a very long time in the works. As the title suggests, the argument draws upon material from fields of study that are not usually in close communication. My ability to speak with any degree of understanding in many of these areas is due to the quality of the instruction that I have received over the years. As a result, I am very pleased to acknowledge first and foremost the contribution that has been made to this work by many of the fine teachers I have had—at McGill: James Tully, Charles Taylor, David Davies; and at Northwestern: Thomas McCarthy, James Johnson, Jeroen Swinkels, Michael Williams, and Jürgen Habermas. The idea of using rational choice theory to evaluate the claims that Habermas makes about the limits of the instrumental conception of rationality is one that I picked up from Jim Johnson. As well, the emphasis on accountability as the key characteristic of social action is something that Tom McCarthy impressed upon me, and that forms the core of his own systematic views.

Many people have read and commented on portions of this work over the years. They include Joel Anderson, Jim Bohman, David Davies, Gordon Finlayson, Jim Johnson, Tom McCarthy, Cheryl Misak, and Kevin Olson. I would like to thank them for their efforts. Also, I am indebted to the graduate students at the University of Toronto who have participated in my various seminars, especially Klaus Jahn, who lent further assistance in the preparation of the final manuscript.

This book incorporates, in revised form, material from previously published articles. I gratefully acknowledge the work of the editors and referees of the journals in which the articles appeared. Sections 2.3 and 2.4 are drawn from

"Is Language a Game?" *Canadian Journal of Philosophy* 26: 1 (Lethbridge: University of Calgary Press, 1996): 1–28; §3.1 takes a bit from "Threats, Promises, and Communicative Action," *European Journal of Philosophy* 3: 3 (Oxford: Blackwell, 1995): 225–241; §3.4 repeats material from "What Is a Validity Claim?" *Philosophy and Social Criticism* 24: 4 (London: Sage, 1998): 23–41; some of §4.1 and §4.4 was published in "The Structure of Normative Control," *Law and Philosophy* 17: 4 (Dordrecht: Kluwer Academic, 1998): 419–442; some of §6.2 in "The Problem of Foundationalism in Habermas's Discourse Ethics," *Philosophy and Social Criticism* 24: 4 (London: Sage, 1998): 23–41; and the argument in §§7.2–7.3 appears as "A Pragmatist Theory of Convergence" in *Pragmatism: Canadian Journal of Philosophy Supplementary Volume*, ed. C. J. Misak (Lethbridge: University of Calgary Press, 1999). My thanks to Blackwell Publishers, Sage Publishers, Kluwer Academic Publishers, and the University of Calgary Press for their kind permission to use this material.

Finally, I would like to thank the University of Toronto Connaught Fund, along with the Social Sciences and Humanities Research Council of Canada for financial assistance at various points in the development and execution of this project.

Contents

Introduction

For Kant the moral order "within" was an awesome mystery; for sociologists the moral order "without" is a technical mystery.

—*Harold Garfinkel*

This book is a critical study of Jürgen Habermas's theory of communicative action. Habermas's theory is important because it is the best attempt that has been made so far to produce a *general theory of social action*. It is a conspicuous feature of the social sciences that different disciplines are dominated by different theories about how people act—what psychological states govern their decisions, what incentives they respond to, and how the actions of multiple agents are coordinated. The most extreme disciplinary division has been between economists and sociologists. Economists have tended to emphasize the instrumental dimension of human action—the way that practical deliberation is concerned with discovering the most efficient means to the realization of given ends—whereas sociologists have argued that "social norms," or shared rules of conduct, play an important role in social action by directly motivating agents to conform to specific institutional patterns. Although these two views are not strictly incompatible, there is a significant tension between them. Insofar as agents are really following rules (i.e., respecting deontic constraints), there is no clear sense in which they can be selecting their actions as mere means to some further end. This tension has given rise to well over a century of debate, and although this has generated significant refinement of the positions on both sides, there has been very little narrowing of the gulf between them. In fact, the seeming intractability of this dispute has led

many theorists to reject abstract social theory entirely in favor of local, interpretive studies.[1]

From Habermas's perspective, the reason that little progress has been made in this debate is that two pieces of the puzzle are missing. First, even though many social scientists have abandoned positivism as a methodological stance, the idea that "value judgments" or "moral questions" are rationally undecidable is still widely accepted. As a consequence, most social theorists simply assume that any agent who acts on the basis of a moral principle, or a social norm, is not rationally justified in doing so. This is what underlies the widespread tendency among social theorists to assume that instrumental action is the only form of rational action, and that norm-governed action must have some kind of nonrational source, such as conditioning, socialization, or habit. This leads to problems, however, because norm-governed action is often highly organized, reflexive, and adaptive.[2] If it is not rational, then the coherence of norm systems is extremely difficult to explain, since it could not be achieved through any specifically intentional process. This makes it tempting to abandon the action frame of reference and supply a purely functionalist explanations of these norms. This strategy has generated a lengthy and increasingly sterile debate about the merits of "micro" versus "macro" explanations, and the status of "methodological individualism."[3]

The most obvious problem with this debate, from a philosophical perspective, is that the assumptions about rationality that are accepted by all parties presuppose an extremely controversial form of moral skepticism. The traditional reason for thinking that normative commitments are irrational, or unjustifiable, depends upon a rather specific conception of rationality and justifiability known as *foundationalism*.[4] This conception of rationality, however, is one that has become increasingly discredited in philosophical circles, largely as a consequence of the so-called linguistic turn.[5] The impact of this change has been felt in debates over scientific methodology, but has still had little impact on the working theories of social scientists. Habermas is one of the first to bring a nonfoundationalist conception of rationality to the task of understanding the logic of social action. One of the consequences of this view of rationality is that he is not inclined to treat moral action as irrational. This means that he can explain conformity to social norms as a straightforward exercise of *rational choice*—except that "rational" is now to be understood in a noninstrumental sense. This allows him to explain the coherence and adaptability of norm-governed systems by appealing to

cognitive resources that agents themselves can deploy in managing their interactions.

The second major piece of the puzzle that has been missing is an account of language. It may come as a surprise to some to discover that the instrumental model of social interaction most often used in the social sciences explicitly excludes any kind of linguistic communication among agents.[6] Given the absolute centrality of language for all forms of social life, it is hard to imagine that a theory of action that excludes communication could have more than limited applicability. A general theory of rational action must give some account of all rational activities—not just consumption and voting decisions, but also such paradigmatically rational activities as doing arithmetic, compiling data in a lab, debating economic policy, and even reading a book on rational choice theory.

This is a problem not just for instrumental theories of action. Theories that incorporate noninstrumental forms of action have generally not been any more successful at accommodating language and communication. Symbolic interactionism, for instance (despite the title), has remained focused on the way that specific actions acquire symbolic properties, and has never been able to explain how propositionally differentiated speech could emerge out of social action.[7] As a result, it provides no systematic account of our most powerful linguistic capacity, namely, the ability to use compositional semantic resources to *represent* possible states of affairs. Without such an account, it is impossible to explain how we use language to make requests, announce intentions, and so forth.

Habermas overcomes this deficit by incorporating *speech act theory* into his model of social action.[8] Thus his theory of action includes a full-scale theory of linguistic action. As a result, the cognitive resources that are simply presupposed by most action theories form an integral part of Habermas's own. For instance, most action theories presuppose that agents come equipped with all sorts of intentional states—beliefs, desires, preferences, values, and so on. Since all these intentional states are propositional attitudes, and since language is a social phenomenon, it would seem that the theory of social action should have the capacity to explain how these states come about (or at least how they acquire their *content*). But any action theory that starts out by *presupposing* these states will be structurally incapable of providing such an account. Habermas's theory of communicative action, because it incorporates a theory of meaning, is not vulnerable to this difficulty. It thus stands as an

exemplar of how a philosophically sophisticated theory of action should be constructed.

The first half of this book is concerned with evaluating Habermas's action theory. At various points in its development, Habermas presents this theory as a part of a general critique of instrumental rationality (this is the central line of continuity between his work and the "Frankfurt school" tradition of critical social science). However, despite being pitched as a critique of instrumental rationality, Habermas's work is informed by only a somewhat vague grasp of the details of the instrumental view. My goal in the present work is to bring Habermas's theory into dialogue with the most sophisticated articulation of the instrumental conception of practical rationality: Bayesian decision and game theory. More specifically, I try to show that Habermas's central criticism of the instrumental view is broadly consistent with limitative results that have been obtained in game-theoretic analysis. The weaknesses that Habermas points to from "outside" the theory have also been pointed out from the "inside"—often more perspicuously. Thus critical theory and game theory, as James Johnson has observed, "converge in improbable but potentially productive ways."[9]

I begin in chapter 1 by presenting a basic outline of Habermas's theory of action. One of the major problems that arises from Habermas's lack of precision with respect to the instrumental model is that his theory has had little impact among those who are not already interested in finding an alternative to the instrumental view. By contrast, Donald Davidson's critique of decision theory, which is in many ways quite similar to Habermas's, has had a much greater effect.[10] This is despite the fact that Davidson's overall action theory is in some respects much less sophisticated than Habermas's (e.g., Davidson does not even try to generalize the action theory to social contexts). The difference is that Davidson gets the decision theory right, and so knows how to make his arguments stick. Since there are broad structural similarities between Davidson's view and Habermas's—particularly on the relationship between rationality and interpretation—I draw upon Davidson's work at various points in order to help situate Habermas's views.

In chapter 2, I present a general outline of decision and game theory, followed by an analysis of Habermas's major criticism of this model. The key idea in Habermas's theory of communicative action is that speech acts cannot be planned or executed with entirely strategic intent. To establish this claim, however, Habermas draws upon some rather controversial work in the philos-

ophy of language—in particular, J. L. Austin's distinction between illocutionary and perlocutionary components of speech acts.[11] My major task in this chapter is to show that Habermas's claim can be established without appealing to any of these problematic claims, but through an entirely game-theoretic analysis of the effects of communication systems on strategic interaction. I develop this argument by showing that interactions among instrumentally rational agents cannot confer semantic content upon any of their actions, and so language could not arise out of strategic interaction. On the other hand, if agents are simply assumed to have primitive linguistic competencies, then they will be incapable of solving social interaction problems using strictly instrumental resources. The upshot of these two arguments is that recent game-theoretic research lends considerable support to Habermas's claim that speech acts cannot be instrumentally rational.

This discussion takes care of the "critical" aspect of Habermas's action theory. However, unlike many critics of the instrumental conception of rationality, Habermas does not forget the old saying that "you can't beat something with nothing." Perhaps the more exciting feature of his work is that instead of just criticizing the instrumental view, as so many others have done, Habermas attempts to present a concrete alternative. This is what I examine in chapter 3. The key here again is Habermas's claim that speech acts are not instrumentally rational. Since there must be some sense in which language use is rational, analyzing the "logic" of speech may help to reveal the "logic" of linguistic action. Here the connection between rationality and interpretation becomes important. To interpret each others' utterances, we must assume that we are at least minimally rational. This means that we can look at our practices of interpretation—how we confer semantic content upon expressions—in order to discover what conception of practical rationality governs our linguistic practices. In a sense, the best theory of meaning imposes a particular conception of practical rationality upon us. This argumentation strategy is what Habermas employs to develop his alternative, "communicative" conception of practical rationality.

Although I endorse this general argumentation strategy, I have some disagreements with the specific theory of meaning that Habermas develops. At the end of the third chapter, I criticize Habermas's attempt to analyze speech acts using three different "validity claims" and outline an alternative to this view. However, since the theory of meaning imposes a theory of practical rationality, the changes that I suggest at the level of semantic theory have

ramifications for the entire theory of action. In chapter 4, I present the broad outlines of a multidimensional theory of rational action that includes "norm-governed" action as a specific type. This theory is similar to Habermas's, but adopts a different explanatory strategy when it comes to accounting for certain features of social order. The result is a model of rational action that contains roughly the same components as Habermas's, but arranges them in a different way.

The second half of this book is dedicated to a discussion of the more philosophical dimension of Habermas's conception of practical rationality. Because he does not approach the theory of action with a prior commitment to moral noncognitivism, Habermas is able to dispel much of the mystery that has surrounded the sociological conception of "norm-governed" action. However, this perspective will only be persuasive if he is able to give some account of moral reasoning that redeems his cognitivist intuitions. The attraction of Habermas's position, in this respect, is that instead of trying to provide an autonomous, "foundational" conception of moral reasoning that would ground the action theory, he imports elements of the action theory into his philosophical ethics, in order to explicate the *practice* of moral argumentation. Thus he draws upon the action-theoretic concept of a "social norm" in order to dispel certain classic problems in moral philosophy. So at the same time that suspending the noncognitivism postulate enables him to develop a more powerful theory of social action, this enhanced action theory is itself deployed in order to relieve some of the skeptical pressures that have traditionally favored moral noncognitivism.

The significance of Habermas's work in moral philosophy has more to do with the way that he situates the major problems than in the specific proposals that he has for resolving them. His major achievement in this respect is to have recognized that the skeptical problem that has motivated many popular forms of moral noncognitivism (often expressed by the question "why be moral?") is equivalent to the technical problem in theoretical sociology referred to as the "problem of order"—or more generally, that the solution to one of these problems would constitute a solution to the other.[12]

The "why be moral?" question is normally taken as a shorthand way of referring to a certain type of skeptical challenge to morality.[13] The question is usually not meant literally. When Hume asked why he should not prefer the destruction of the world to the scratching of his finger,[14] it would always have been possible to provide a straightforward response, for example, "The