The Grammar of Society

The Nature and Dynamics of Social Norms

In *The Grammar of Society*, Cristina Bicchieri examines social norms, such as fairness, cooperation, and reciprocity, in an effort to understand their nature and dynamics, the expectations that they generate, and how they evolve and change. Drawing on several intellectual traditions and methods, including those of social psychology, experimental economics, and evolutionary game theory, Bicchieri provides an integrated account of how social norms emerge, why and when we follow them, and the situations in which we are most likely to focus on relevant norms. Examining the existence and survival of inefficient norms, she demonstrates how norms evolve in ways that depend on the psychological dispositions of the individual and how such dispositions may impair collective welfare. By contrast, she also shows how certain psychological propensities may naturally lead individuals to evolve fairness norms that closely resemble those we follow in modern societies.

Cristina Bicchieri is the Carol and Michael Lowenstein Term Chair and Professor of philosophy at the University of Pennsylvania, where she is also the Director of the Philosophy, Politics, and Economics Program. The author of many articles and books, including *Rationality and Coordination, The Logic of Strategy*, *The Dynamics of Norms*, and *Knowledge, Belief and Strategic Interaction*, she has been a Fellow at the Wissenschaftskolleg zu Berlin, the Swedish Collegium for Advanced Study in the Social Sciences, the London School of Economics (Leverhulme Trust), and the Institute for Advanced Studies at the University of Jerusalem.

The Grammar of Society

The Nature and Dynamics of Social Norms

CRISTINA BICCHIERI
University of Pennsylvania

CAMBRIDGE UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by
Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521573726

First published 2006

A catalogue record for this publication is available from the British Library

Library of Congress Cataloguing in Publication data

Bicchieri, Cristina.
The grammar of society : the nature and dynamics of social norms /
Cristina Bicchieri.
p. cm.
Includes bibliographical references and index.
ISBN 0-521-57372-6 (hardcover : alk. paper) – ISBN 0-521-57490-0 (pbk. : alk. paper)
1. Social norms. I. Title.
HM676.B53 2006
306–dc22 2005006200

ISBN 978-0-521-57372-6 Hardback
ISBN 978-0-521-57490-7 Paperback

To my parents Vanna and Ettore
and my daughter Tatiana, who
taught me much about
following and breaking rules

Contents

Preface

I do not remember when my interest in social norms began, but the subject has been a long-standing source of curiosity and frustration for me. As a stranger living for many years in foreign countries, I have had to constantly negotiate the meaning of rules and practices that more often than not I did not fully understand, the subtleties of a social language that was not my mother tongue. Norms are the language a society speaks, the embodiment of its values and collective desires, the secure guide in the uncertain lands we all traverse, the common practices that hold human groups together. The norms I am talking about are not written and codified; you cannot find them in books or be explicitly told about them at the outset of your immersion in a foreign culture. We learn such rules and practices by observing others and solidify our grasp through a long process of trial and error. I call social norms *the grammar of society* because, like a collection of linguistic rules that are implicit in a language and define it, social norms are implicit in the operations of a society and make it what it is. Like a grammar, a system of norms specifies what is acceptable and what is not in a social group. And analogously to a grammar, a system of norms is not the product of human design and planning.

My fascination with norms has thus been both personal and intellectual. I am always surprised to realize that norms are supported by and in some sense consist of a cluster of self-fulfilling expectations. If people believe that a sufficiently large number of others uphold a given norm, then, under the right conditions, they will conform to it. A norm's destiny is strictly connected to the dynamics of such expectations; a change in expectations may lead to a dramatic decline in norm compliance and

to the eventual demise of the norm itself. How such expectations are formed, where they come from, is one of the themes I address in this book. My frustration has similarly had both very personal and intellectual facets – personal because learning a society's norms as an adult is far less natural and effortless than when you are born into a given culture, and intellectual because much of what is written about norms does not seem to capture what I consider to be their essential features. This book is an answer to my deep-rooted questions about the nature of norms, how they can emerge and thrive or decay, and what compels people to follow them.

The social norms I am talking about are not the formal, prescriptive or proscriptive rules designed, imposed, and enforced by an exogenous authority through the administration of selective incentives. I rather discuss informal norms that emerge through the decentralized interaction of agents within a collective and are not imposed or designed by an authority. Social norms can spontaneously develop from the interactions of individuals who did not plan or design them, as can conventions and descriptive norms. All three are social constructs that have a life simply because enough people believe they exist and act accordingly. To distinguish between these three very different social constructs, I focus in Chapter 1 on the kinds of situations in which they are likely to emerge, as well as on the types of expectations and preferences that support them. Descriptive norms such as fashions and fads, for example, arise in contexts in which people desire to coordinate with (or imitate) others and prefer to do what others do on the condition that they expect a sufficient number of people to act in a certain way. A 'sufficient number' may be just one person, as in the case of a celebrity we want to imitate, or the number may vary from person to person, depending on how cautious one is in assessing the threshold at which to take action. Conventions are descriptive norms that have endured the test of time. If one's main objective is to coordinate with others, and the right mutual expectations are present, people will follow whatever convention is in place. Social norms, on the contrary, are not there to solve a coordination problem. The kinds of situations to which social norms most often apply are those in which there is a tension between individual and collective gains. Pro-social norms of fairness, reciprocity, cooperation, and the like exist precisely because it might not be in the individual's immediate self-interest to behave in a socially beneficial way. This does not mean we follow such norms only when coerced to do so. Granted, some people need incentives in the form of the expectation of rewards and punishments to be induced to comply.

Others instead obey a norm just because they recognize the legitimacy of others' expectations that they will follow the norm. My definition of what it takes for a social norm to exist and be followed takes into account the fact that there are different types of people. All have conditional preferences for conformity, and all need to believe that enough people are obeying the norm to make it worthwhile to conform. What makes people different is the nature of their normative expectations: Some just need to believe that enough other people expect them to conform, whereas others need to believe that others are also prepared to punish their transgressions. In both cases, I stress that preference for conformity is *conditional*. If expectations change, so does conforming behavior. I maintain that norms are never the solution of an original coordination game. However, once a norm is in place, it will *transform* the original game into a new coordination game, at least for those who believe that the norm is in fact followed. In the new game the choice is to follow the standing norm or 'defect' and thus revert to the original game. This choice depends on what we expect others to do. These expectations may be grounded in our knowledge of past behavior of the people we interact with, but more often than not we do not have such personal knowledge of our parties. Where our expectations come from and what grounds them is the theme of Chapter 2.

Because the important question is not whether norms affect behavior, but when, how, and to what degree, in Chapter 2 I show under which conditions the beliefs and preferences that support a norm are activated as the result of the interpretation of specific cues, the categorization of the situation based on those cues, and the consequent activation of appropriate scripts. A situation can be interpreted and categorized in several ways, with very different consequences for norm compliance. An observed exchange, for example, can be perceived as a market interaction, an instance of gift giving, or an act of bribing. Depending on how we categorize it, our expectations, predictions, and emotional responses will be very different. Categorizing an exchange as an instance of gift giving will activate a script that specifies, among other things, roles and possible action sequences. Norms, I argue, are embedded into scripts, the rudimentary theories about social roles and situations that guide us in interpreting social interactions, forming expectations and predictions, assessing intentions, and making causal attributions. Once a script has been activated, the corresponding beliefs, preferences, and behavioral rules (norms) are prompted. The expectations and preferences that determine our choices are thus the result of the activation of collectively

shared scripts that are general enough to subsume a wide variety of situations.

The only systematic evidence presently available about which cues make people focus on particular norms are the results of experiments on Ultimatum, Dictator, Trust, and Social Dilemma games. Though the experiments I discuss in Chapters 3 and 4 were not meant to test hypotheses about norms, their results are consistent with a theory of script activation. Furthermore, I show that some behavioral inconsistencies that have baffled investigators can become comprehensible in light of the view of norms I am proposing. In Chapter 3, I consider experimental Ultimatum and Dictator games and contrast the social preference models that have been proposed to explain the results with my own norm-based utility function. I hope to convince the reader that such a utility function is more general than many of those that have been proposed, and that it makes interesting, testable predictions about how manipulating subjects' expectations may induce, or eliminate, conformity to a norm. Chapter 4 examines social dilemma experiments and the surprising results obtained by allowing pre-play communication among the players. When subjects are permitted to communicate about the experiment, even if for a very brief time, we observe almost universal cooperation. A favored explanation is that communication creates a social identity, an *esprit de corps* that would induce a deep change in preferences. I examine the merits of the social identity hypothesis but argue that the available data do not support it. Instead, they support an explanation in terms of social norms. Communication is particularly successful when people make promises to each other, and even if the one-shot nature of the interaction should make such promises no more than cheap talk, it is sufficient to yield scripts (and norms) that support cooperative behavior.

In the last two chapters I look at how a norm might emerge in a situation in which there is none – individuals may, nonetheless, believe a norm exists and actively try to conform to it. A common assumption many people tend to make is that if a norm emerges, then it must be socially advantageous or efficient. In Chapter 5, I show that the dynamics of norm formation may be such that a bad descriptive norm or a bad convention can easily come about if certain conditions are present. And the transformation of such a bad convention into a poor social norm is always possible. The most common condition in which a bad norm is likely to occur is one in which individuals are in a state of *pluralistic ignorance*. When there is an incentive to conform to what other people do, there is no transparent communication, and individuals have a tendency to believe

that what they observe others doing reflects their true preferences, then it is likely that the collective outcome will be something most participants did not want and may even keenly dislike. Again, this is a particularly powerful example of the role collective beliefs play in generating social institutions that may turn out to be far from efficient or socially beneficial.

In Chapter 6, I look at how a social norm, a norm of fair division in this case, can evolve from the interactions of agents who believe a norm exists but have no idea what it is. I assume that agents care in varying degrees about norms, and that they are trying to learn what the shared norm is, because they wrongly believe there must be one. This model is quite different from the traditional evolutionary models we find in the literature on the evolution of norms, especially because it starts from specific psychological assumptions about individual dispositions, assumptions that are in fact well supported by psychological research. The interesting result is that individuals endowed with such dispositions who interact with each other and are capable of learning and revising their strategies according to a best-reponse dynamics will indeed generate a norm of fair division. Such a norm is very close to the modal and median offers we observe in experimental Ultimatum games. Much work remains to be done about how certain dispositions to recognize and follow norms have evolved, and why. What I want to show is that norms can endogenously emerge from the interactions of individuals who share such dispositions, and I hope I have convinced my readers that this is a real possibility.

Acknowledgments

This book took a very long time to write, and I cannot hope to recall all the colleagues and students with whom, at various times, I have discussed some of the topics it addresses. Brian Skyrms offered constant encouragement and intellectually stimulating exchanges over the years. Robyn Dawes's friendship and intellectual companionship has given me more than I will ever be able to acknowledge. His criticism was always constructive, and he guided my first steps in the maze of experimental research. Colleagues at Carnegie Mellon, the Center for Rationality and Interactive Decision Theory at the University of Jerusalem, and the Swedish Collegium for Advanced Study in the Social Sciences in Uppsala have read earlier versions of many chapters and have provided advice, encouragement, and constructive criticism. Part of the book was written during the year I spent at the Wissenschaftskolleg in Berlin. The norms seminar we had every week was a fun, stimulating, and challenging arena in which to test some of my ideas: Ernst Fehr, Robert Boyd, Joe Heinrich, Alex Kacelnik, Michael Kosfeld, and (too briefly) Shlomo Shwartz and Catherine Eckel were wonderful partners in the quest for understanding why we (and animals) behave as we do.

My students had an important role to play, because many of them were exposed to several versions of the manuscript, and their thoughtful comments helped make it what it is. Erin Krupka edited some of the chapters and always offered intelligent advice, as did Giacomo Sillari and Ryan Muldoon, who read and reread the whole manuscript and gave me precious suggestions. Daniel Neill and Jiji Zhang were wonderful helpers, always ready to pick a formal challenge. Jason Dana's knowledge of the experimental literature was a boon, especially with respect to the work

done on fairness. Finally, several friends and colleagues read and commented on different versions of specific chapters of the book. A special thanks goes to Jason Alexander, Peter Vandershraaf, Rachel Croson, John Duffy, Yoshi Fukui, Massimo Bigliardo, Matteo Motterlini, Bill Keech, Roberto Weber, and Peyton Young.

Chapters 4 and 5 are based on previous articles. Chapter 4 draws from "Covenants Without Swords: Group Identity, Norms, and Communication in Social Dilemmas," *Rationality and Society* 14(2): 192–228, 2002. Chapter 5 contains part of the material that appeared in "The Great Illusion: Ignorance, Informational Cascades and the Persistence of Unpopular Norms" (with Y. Fukui), *Business Ethics Quarterly* 9: 127–155, 1999. I wish to thank those journals for allowing me to publish some of the material. Some of the material of Chapter 6 will appear in "The Evolution of Fairness," an article jointly written with Jason Alexander.