

Module 2 Unit 3

This is a **REQUIRED READING**.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. New York: Cambridge University Press. [preface and pp. 1 – 3]. [8 p.]

The Grammar of Society

The Nature and Dynamics of Social Norms

CRISTINA BICCHIERI

University of Pennsylvania



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-response 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.

The Rules We Live By

Introduction

Despite the ubiquitous reference to the concept of social norms in the social sciences, there is no consensus about the power of social norms to direct human action. For some, norms have a central and regular influence on human behavior, while for others, the concept is too vague, and the evidence we have about norm compliance is too contradictory to support the claim that they appreciably affect behavior. Those who doubt that norms have a behavior-guiding force argue that human behavior only occasionally conforms with the dominant social norms. If the same norms are in place when behavior is norm-consistent as when it is norm inconsistent, why should we believe that norms mediated any of it?

Much of the discussion about the power norms have to affect behavior arises from a confusion about what is meant by 'norm.' A norm can be formal or informal, personal or collective, descriptive of what most people do, or prescriptive of behavior. In the same social setting, conformity to these different kinds of norms stems from a variety of motivations and produces distinct, sometimes even opposing, behavioral patterns. Take for example a culture in which many individuals have strong personal norms that prohibit corrupt practices and in which there are legal norms against bribing public officers, yet bribing is widespread and tolerated. Suppose we were able to independently assess whether an individual has a personal norm against corruption. Can we predict whether a person, who we know condemns corruption, will bribe a public officer when given a chance? Probably not, but we could come closer to a good prediction if we knew certain factors and cues are present in this situation and have

an influence on the decision. The theories of norms we have inherited, mainly from sociology, offer little help, because they did not develop an understanding of the conditions under which individuals are likely to follow a norm or, when several norms may apply, what makes one of them focal.

A first step in the direction of a deeper understanding of what motivates us to follow a norm is to clarify what we mean by a social norm. 'Norm' is a term used to refer to a variety of behaviors, and accompanying expectations. These should not be lumped together, on pain of missing some important features that are of great help in understanding phenomena such as variance in norm compliance. Inconsistent conformity, for example, is to be expected with certain types of norms, but not with others. In this chapter I put forth a 'constructivist' theory of norms, one that explains norms in terms of the expectations and preferences of those who follow them. My view is that the very existence of a social norm depends on a sufficient number of people believing that it exists and pertains to a given type of situation, and expecting that enough other people are following it in those kinds of situations. Given the right kind of expectations, people will have conditional preferences for obeying a norm, meaning that preferences will be conditional on having expectations about other people's conformity. Such expectations and preferences will result in collective behaviors that further confirm the existence of the norm in the eyes of its followers.

Expectations and conditional preferences are the building blocks of several social constructs, though, not just social norms. *Descriptive norms* such as fashions and fads are also based on expectations of conformity and conditional preferences, and so are *conventions*, such as signaling systems, rules of etiquette, and traffic rules. In both cases, the preference for conformity does not clash with self-interest, especially if we define it in purely material terms.¹ One can model descriptive norms and conventions as solutions to coordination games. Such games capture the structure of situations where there exist several possible equilibria and, although we might like one of them best, what we most want is to coordinate with others on *any* equilibrium; hence we act in conformity to what we expect others to do. Descriptive norms and conventions are thus representable as equilibria of original coordination games. *Social norms*, on the contrary, often go against narrow self-interest, as when we are

¹ What one most prefers in these cases is to 'do as others do,' or to coordinate with others' choices.

required to cooperate, reciprocate, act fairly, or do anything that may involve some material cost or the forgoing of some benefit. The kinds of problems that social norms are meant to solve differ from the coordination problems that conventions and descriptive norms 'solve.' We need social norms in all those situations in which there is conflict of interest but also a potential for joint gain. The games that social norms solve are called mixed-motive games.² Such mixed-motive games are not games of coordination to start with, but social norms, as I shall argue, *transform* mixed-motive games into coordination ones. This transformation, however, hinges on each individual expecting enough other people to follow the norm, too. If this expectation is violated, an individual will revert to playing the original game and to behaving 'selfishly.' This chapter thus starts with a precise definition of social norms and only later considers what differentiates such norms from descriptive norms and conventions. Because all three are based on expectations and conditional preferences, I pay special attention to the nature of expectations (empirical and/or normative) that support each construct.

The definition of social norm I am proposing should be taken as a *rational reconstruction* of what a social norm is, not a faithful descriptive account of the real beliefs and preferences people have or of the way in which they in fact deliberate. Such a reconstruction, however, will have to be reliable in that it must be possible to extract meaningful, testable predictions from it. This is one of the tasks I undertake in Chapters 3 and 4. An important claim I make in this chapter is that the belief/desire model of choice that is the core of my rational reconstruction of social norms does not commit us to avow that we always engage in conscious deliberation to decide whether to follow a norm. We may follow a norm automatically and thoughtlessly and yet still be able to explain our action in terms of beliefs and desires.

The simplistic, common view that we conform to norms either because of external sanctions or because they have been internalized flies in the face of much evidence that people sometimes obey norms even in the absence of any obvious incentive structure or personal commitment to what the norm stands for (Cialdini et al. 1990). Many who postulate internal or external incentives as the sole reasons for compliance also maintain compliance is the result of a conscious process of balancing costs

² Well-known examples of mixed-motive games that can be 'solved' (or better, 'transformed') by norms of fairness, reciprocity, promise-keeping, etc., are the Prisoner's Dilemma, the Trust game, and Ultimatum games.