

Review of Cristina Bicchieri's *Norms in the wild: how to diagnose, measure, and change social norms*. Oxford: Oxford University Press, 2017, xviii + 239 pp.

FRANCESCO GUALA

Università degli Studi di Milano

In her 2012 Descartes Lectures at Tilburg University, Cristina Bicchieri addressed three questions that have been central in her research for the last twenty years: What are social norms? How can they be measured? And how can we change them? The monograph under review is a descendant of those lectures. Conceived originally as a standard academic volume—with commentaries and replies—the nature of the book changed significantly during its gestation. Since 2010 Bicchieri had been involved in a project aimed at training social workers. The project, led by UNICEF in collaboration with various NGOs, aimed at tackling practices such as breastfeeding prohibitions, female genital mutilation, child marriage, and open defecation—which all had proven to be remarkably resistant to traditional policy interventions. Thanks to this project, Bicchieri had the chance to do what philosophers rarely do, namely, to make her theory practically relevant.

Rather than a classic academic book of philosophy and social science, *Norms in the wild* is a partly theoretical, partly practical manual for policy-makers and social workers interested in changing people's behaviour through the manipulation of norms. The book is organised in five chapters: "Diagnosing norms", "Measuring norms", "Norm change", "Tools for change", and "Trendsetters". The style is direct and simple, with frequent repetitions and summaries ('take-home messages') at the end of sections, to facilitate comprehension and recollection. Footnotes are kept to a minimum, the bibliography is essential, and—importantly—Bicchieri never engages in conceptual or empirical disputes with alternative theoretical frameworks or interpretations of the evidence. Her task here is not to convince researchers, but practitioners and, presumably, politicians.

If you are unfamiliar with Bicchieri's earlier work and with related research on social norms, this is a good point to start. The monograph is well-organized, clear, delivers complex ideas in the simplest manner,

and offers an overview of how these ideas can be applied to concrete problems. *Norms in the wild* is not theoretically innocent, to be sure. But new ideas are not highlighted as such, nor are they discussed in the context of recent research. As a consequence, a sophisticated reader must do most of the work of reconstructing the existing debate herself, as well as placing the development of Bicchieri's theory in its context.

I. WHAT ARE NORMS?

Bicchieri has tried to give an answer to this question for more than two decades. The key elements can be found already in *Rationality and coordination* (1993), and have been developed in later work with slight modifications. So, a curious reader will want to know whether *Norms in the wild* offers anything new from this respect.

Bicchieri's interest in norms was originally prompted by the puzzle of cooperation. How do we explain the fact that many people—in experiments as well as, arguably, in real life—engage in cooperation in situations in which it is a dominated strategy? A standard way to tackle this question, both in the philosophical and in the scientific literature, has been to tinker with individual payoffs. The idea is that people may be motivated by factors that go beyond the narrow selfish preferences of standard economic models. Bicchieri's original twist has been to ignore people's moral concerns—their taste for equality, for example, their altruism, or their desire to reciprocate 'nice' and 'nasty' actions—and to focus instead on social norms. A norm is essentially a shared preference to follow a behavioural rule, conditional on the fact that other people follow that rule and expect others to do it. Her 1993 definition goes like this:

Let R be a behavioural regularity in population P . Then more generally, R is a social norm iff R depends on the beliefs and preferences of the members of P in the following way:

- (1) Almost every member of P prefers to conform to R on the condition (and only on the condition) that almost everyone else conforms, too.
- (2) Almost every member of P believes that almost every other member of P conforms to R (p. 232).

Bicchieri's idea was developed further in her second, very successful monograph, *The grammar of society* (2006). The *Grammar* not only contained a richer theoretical account of social norms, but also backed

it up with a wealth of evidence drawn from social psychology, experimental and behavioural economics (see Guala 2007).

In the *Grammar* we find the following, more sophisticated account:

R is a social norm in a population *P* if there exists a sufficiently large subset of *P* such that, for each individual belonging to this subset:

- (1) *i* knows that a rule *R* exists and applies to situations of type *S*;
- (2) *i* prefers to conform to *R* in situations of type *S* on the condition that:
 - (a) *i* believes that a sufficiently large subset of *P* conforms to *R* in situations of type *S*, and either
 - (b) *i* believes that a sufficiently large subset of *P* expects *i* to conform to *R* in situations of type *S*, or
 - (b') *i* believes that a sufficiently large subset of *P* expects *i* to conform to *R* in situations of type *S*, prefers *i* to conform, and may sanction behaviour (from Bicchieri 2006, p. 11, with a few slight modifications to simplify the notation).

Condition 2(a) is called “Empirical expectations”, while 2(b) is “Normative expectations” and 2(b') is “Normative expectations with sanctions”. These requirements distinguish Bicchieri’s theory of norms from every significant account proposed earlier, making individual preferences conditional on a set of connected beliefs or expectations.¹ A distinctive feature of the 1993 formulation is that the beliefs in 2 are second-order *empirical* expectations. At first sight, the formal definition proposed in the *Grammar* seems to replicate this approach, except that in 2(b') the beliefs are backed up by sanctions. A couple of pages later, however, Bicchieri asks:

What sort of belief is this? On the one hand, it might just be an empirical belief. If I have consistently followed *R* in situations of type *S* in the past, people may reasonably infer that, *ceteris paribus*, I will do the same in the future, and that is what I believe. On the other hand, it might be a normative belief: I believe a sufficiently

¹ Regarding 2(a), I still think it is odd to say that John’s preference to conform is conditional on his *belief* that others conform: unless John is an idealist philosopher, surely, he is more concerned about what Jane will do, than about his own belief about what Jane will do. Of course, John’s action—his decision to conform or not—will depend on his beliefs about Jane’s behaviour. But his *preference*, rationally, should not (see Guala 2007).

large number of people think that I have an obligation to conform to *R* in the appropriate circumstances (2006, p. 15).

This annotation changes the nature of the theory significantly: while her earlier account attempted to analyse social norms in terms of concepts that did not include any normative element, Bicchieri's 2006 account allows norms or normative terms to appear on both sides of the equation. Normative expectations are not second-order empirical beliefs anymore. There may be an explicit 'ought' in condition 2(b), which is not further analysed.

The normative character of expectations is further emphasised in *Norms in the wild*, and made more explicit. Here is Bicchieri's new definition of social norm:

A social norm is a rule of behaviour such that individuals prefer to conform to it on condition that they believe that (a) most people in their reference network conform to it (empirical expectation) and (b) that most people in their reference network believe they ought to conform to it (normative expectation) (p. 35, italics in original).

Another way to put it is to say that *Rationality and coordination* offered a reductive account of social norms, while *The grammar of society* and (more explicitly) *Norms in the wild* do not. The non-reductive account is not circular—norms are not analysed in terms of norms, but rather in terms of individual normative beliefs. Nonetheless, the notion of normative belief is not explained further. One possibility is to say that a normative expectation is simply an empirical expectation augmented by a propensity to punish deviant behaviour (along the lines of condition 2(b') in the *Grammar*). Another solution is to say that normative expectations are constituted by various propensities that may include, say, mental representations (the propensity to represent future behaviours), verbal reactions (the propensity to praise or reproach), emotions (the propensity to be surprised or annoyed), and so on. In other words, the sort of 'reactive attitudes', à la Strawson, that we use on a daily basis to regulate each other's behaviour.

Following the latter approach would make Bicchieri's theory rather similar to its main current rival, the account of norms as clusters of normative attitudes proposed by Geoffrey Brennan et al. (2010). Brennan and co-authors explicitly offer a non-reductive account, and defend it

using various ingenious arguments (aimed in particular at Bicchieri's theory). Although I am not convinced by some of these arguments, I wonder whether Bicchieri has anything to say about them. It is a pity that alternative accounts of norms are not discussed in *Norms in the wild*—or, as far as I know, anywhere else in Bicchieri's writings.

Another issue that remains unexplored is the motivational basis of norm compliance. When beliefs are backed up by the threat of sanctions (condition 2(b') in the *Grammar*), the instrumental motive to comply is pretty obvious. But what about condition 2(b), or condition (b) in *Norms in the wild*? Why should I be motivated to do something, merely because someone believes that I ought to do it?

One suggestion is that we might have an innate propensity to care about other people's expectations.² Our failure to comply with norms and conventions is a potential source of unpredictability, and being unpredictable hinders one's capacity to engage in cooperative tasks. This, in turn, may reduce one's fitness. So, it would not be surprising, from an evolutionary psychology perspective, if we were endowed with a propensity to fulfil others' expectations. But then we should also expect people to use a variety of ways to make such expectations salient by means of reproach, punishment, and the like. Bicchieri seems to follow this line of reasoning:

the social pressure to conform, expressed in the social expectation that one ought to conform, is a powerful motivator. [...] If others believe one *ought* to conform, the reaction to non conformity may range from slight displeasure to active or even extreme punishment (2017, pp. 34-35, italics in original).

But again, we are spared the details, and the causes of our motivation to fulfil others' expectations remain largely opaque.

II. MEASURING NORMS

A good deal of chapter two ("Measuring norms") is devoted to illustrating experimental and non-experimental techniques to detect and quantify the impact of social norms. Bicchieri has been actively engaged in this research, designing and running laboratory experiments with various collaborators during the last decade. Laboratory experiments, however, are very useful to test causal hypotheses, less so for policy-

² See, e.g., Sugden (1998, 2000). Similar considerations support Brennan and Pettit's (2004) theory of esteem.

making. Because *Norms in the wild* is particularly focused on field intervention, Bicchieri spends a great deal of time illustrating alternative techniques for the measurement of social norms—like surveys, questionnaires, and vignettes.

The limitations of these methods are well-known: they are not incentivized, they do not involve the direct manipulation of key variables, and they do not allow the observation of people's real choices. Still, verbal reports help researchers to elicit peoples' conscious beliefs about others' expectations.

Many techniques illustrated by Bicchieri are based on direct questions about 'what would happen if' someone in the community did this or that, or on the visual presentation of little stories in which fictional but realistic characters engage in behaviours that may be considered problematic from a normative point of view. The fact that such stories are couched in familiar contexts and scenarios is clearly an advantage compared to the anodyne choice situations of many laboratory experiments. Among other things, they facilitate the emergence of plausible responses from our largely subconscious behavioural repertoire. Counterfactual scenarios help individual responders to make explicit the content of rules, and the attitudes that others may have with respect to such rules. An important point is that observation and manipulation are mingled here, because the elicitation of beliefs often constitutes the first step of the policy intervention: *thinking* about norms is one of the conditions that make social change possible.

III. A DUAL ACCOUNT

In the *Grammar* and in *Norms in the wild*, Bicchieri is careful to say that her expectation-dependent model of norms is not supposed to give a literal description of the psychological processes that govern norm compliance in ordinary situations. If you see people queuing in front of a desk, you usually do not begin to ponder: 'What do these people expect me to do?', or 'What would their reaction be if I cut in front of the line?'. In most ordinary cases, you simply apply automatically the rules that apply in the circumstances. Your actions are governed by 'scripts':

Scripts are essentially prescriptive sequences of actions of varying levels of specificity that people automatically engage in (and are expected to engage in) while in particular situations (p. 132).

This account of rule-following is intuitively correct. It is consistent not only with the phenomenology of ordinary action, but also with the bounded rationality, heuristics, and dual-system theories that have come to dominate the cognitive and behavioural sciences over the last thirty years. The interesting issue is what relationship there may be between automatic script-guided behaviour, on the one hand, and the expectation-dependent preferences that are at the core of Bicchieri's definition of norms, on the other.

Bicchieri's writing tends to become metaphorical at this point:

Norms are embedded into scripts because scripts contain empirical and normative expectations, and violations of scripts typically elicit negative emotions and remedial actions (p. 132).

We have beliefs and can refer to them when we want to, but the majority of the time they stay dormant—exerting unconscious influence on our behaviour up until the point when their validity is challenged (p. 128).

How does a script “contain” expectations, exactly? What does it mean that a belief remains “dormant”? How can it “exert *unconscious influence* on our behaviour”? These metaphors unfortunately do not contribute much to our understanding of norm-abiding behaviour.

In light of the evidence surveyed by Bicchieri, it seems more appropriate to admit that expectations play a minor role in ordinary cases of compliance with social norms. Scripts, heuristics, cue-driven automatic behaviour probably do all of the job that needs to be done. But this then means that there are *two* accounts of norm-following in Bicchieri's theory—one for ‘ordinary’ and one for ‘extraordinary’ situations, so to speak. This is an important point that Bicchieri does not emphasise enough, in my view. It is important because it brings her theory in line with recent developments in the philosophy of mind, and because it shows in which sense it may provide a rational-choice explanation of norm-driven behaviour.

What picture of the mind is implicit in Bicchieri's theory? Her definition of norms is essentially game-theoretic, and standard game theory moves from the assumption that behaviour is determined by individual beliefs and desires (or preferences). Moreover, it presupposes that people are able to form beliefs about other people's beliefs and desires. People are belief-desire mind readers.

A substantial body of research, however, has recently questioned the belief-desire model of mind reading that economists and philosophers cherish. This body of research points out that belief-desire attributions are used only rarely to predict and explain other people's behaviour. In fact, they seem to be used mainly in anomalous circumstances—that is, when other techniques of interpretation fail.³

Suppose you are sitting in the waiting room of a public office, with a 'no-smoking' sign hanging on the wall in front of you. The main reason why you don't pull out a cigarette, probably, is that the sign says so. You don't need to think about what other people in the room might possibly think if you smoked. But if a lady is lighting a cigarette, you begin to speculate about her beliefs, and about the expectations of by-standers (you try to meet their gaze, trying to detect cues of their normative attitudes). As Bicchieri points out, "we often become consciously aware of our beliefs the moment something unexpected happens and those beliefs are challenged" (p. 128).

The story Bicchieri is offering therefore is a *dual account* of the psychology of social norms. Norm-following on the one hand, and norm-enforcing and norm-transgressing on the other, are likely to be governed by distinct mechanisms.⁴ But her rational choice definition of norms is able to capture only half of the story.

One may argue that rational choice models are explanatory even when the causal factors they mention (normative expectations, in this case) are not activated. Philip Pettit (1995), for example, has argued that rational choice models can explain the *resilience*, as opposed to the *continuation* of behaviour. Explanations of resilience tell us why a certain pattern does not change, or would not change, in spite of changes in the circumstances. Thus, given a certain structure of incentives and beliefs, people may follow a norm without thinking about those incentives and beliefs at all (continuation). But the important point is that, were the situation to become anomalous, the mechanisms mentioned by Bicchieri (expectation-dependent preferences) would bring the behaviour back in line with the content of the norm (resilience). This

³ The origins of this idea can be found in Bruner (1990) and have been developed in various directions in the following literature. See, for example, McGeer (2007) and Hutto (2008).

⁴ To put it concisely, using the standard two-systems terminology (e.g., Kahneman 2011): norm-following is mostly system 1, while norm change, enforcement, and perhaps also transgression engage (mostly) system 2. Pluralist accounts of mind-reading are defended by Andrews (2012) and Fiebich and Coltheart (2015), for example.

would reconcile Bicchieri's formal definition of norms with the dual account that she outlines informally.

IV. REPRESENTING AND INTERVENING

At this point, it is important to recall what the ultimate goal of Bicchieri's theorizing is. Because her fundamental interest is intervention, norm-enforcement and norm-evasion are salient in her discussion. And because expectations are crucial both to enforce and to change existing norms, the theory places expectations at centre stage. But if the question is *what norms are*, and *how they influence behaviour*, the complete answer must go along these lines: they are rules mapping behavioural responses onto stereotypical situations, that are automatically triggered and followed in ordinary circumstances (for example, when everyone complies); that activate the reactive attitudes of others when they are violated occasionally, and a thorough examination of other people's empirical and normative beliefs when they are systematically or collectively breached by several people. In the latter two cases, our preference for compliance is truly conditional on expectations, that is, we deliberately modulate our behaviour according to others' psychological and behavioural dispositions. Otherwise, expectations do not play a direct causal role in norm-compliance.

There is a sense in which *Norms in the wild* may be read as a guidebook for the creation of anomalous situations, that is, situations in which people stop following rules in an automatic manner, and begin instead to think critically about the attitudes that other people have—or ought to have—about those rules. Social workers, health practitioners, political activists and leaders can play a crucial role in this process that through exposure, recognition, questioning, and collective discussion may lead eventually to the replacement of harmful norms with beneficial ones. Bicchieri's definition of norms, the measurement techniques that she illustrates, and the policy interventions they prepare the ground for, are different parts of one and the same package. To paraphrase Ian Hacking (1983), Bicchieri's theory is not only (perhaps not mainly) for *representing*, but also (perhaps especially) for *intervening*.

This is what makes *Norms in the wild* interesting and different from most books in the philosophy of social science. While reading it, you will learn about the importance of convincing mothers-in-law that feeding colostrum does not harm babies; you will learn how to use flies to

demonstrate that open defecation leads to food poisoning; you will learn why soap operas emancipate women; you will learn what characteristics an individual must have to become a trendsetter; and many other fascinating tricks to manipulate the norms that govern our behaviour.

Bicchieri must be commended for trying to turn a theoretical insight into a toolbox for policy making. If you are tired of philosophers' unlikely thought experiments, if you are fed up with the pursuit of conceptual analysis for its own sake, or if you are worried that social ontology and the philosophy of social science may have no practical implications, then there is no better reading than this.

REFERENCES

- Andrews, Kristin. 2012. *Do apes read minds?* Cambridge (MA): MIT Press.
- Bicchieri, Cristina. 1993. *Rationality and coordination*. New York: Cambridge University Press.
- Bicchieri, Cristina. 2006. *The grammar of society*. New York: Cambridge University Press.
- Brennan, Geoffrey, Lina Eriksson, Robert E. Goodin, and Nicholas Southwood. 2010. *Explaining norms*. Oxford: Oxford University Press.
- Brennan, Geoffrey, and Philip Pettit. 2004. *The economy of esteem*. Oxford: Oxford University Press.
- Bruner, Jerome. 1990. *Acts of meaning*. Cambridge (MA): Harvard University Press.
- Fiebich, Anika, and Max Coltheart. 2015. Various ways to understand other minds: towards a pluralistic approach to the explanation of social understanding. *Mind & Language*, 30 (3): 235-258.
- Guala, Francesco. 2007. Review of *The grammar of society: the nature and dynamics of social norms* by Cristina Bicchieri. *The British Journal for the Philosophy of Science*, 58 (3): 613-618
- Hacking, Ian. 1983. *Representing and intervening*. Cambridge: Cambridge University Press.
- Hutto, Daniel D. 2008. *Folk psychological narratives*. Cambridge (MA): MIT Press.
- McGeer, Victoria. 2007. The regulative dimension of folk psychology. In *Folk psychology re-assessed*, eds. Daniel D. Hutto and Matthew Ratcliffe. Dordrecht: Springer, 137-156.
- Pettit, Philip. 1995. The virtual reality of *homo economicus*. *The Monist*, 78 (3): 308-329.
- Sugden, Robert. 1998. Normative expectations: the simultaneous evolution of institutions and norms. In *Economics, values, and organization*, eds. Avner Ben-Ner and Louis Putterman. Cambridge: Cambridge University Press, 73-100.
- Sugden, Robert. 2000. The motivating power of expectations. In *Rationality, rules, and structure*, eds. Julian Nida-Rümelin and Wolfgang Spohn. Dordrecht: Springer, 103-129.

Francesco Guala is a philosopher and experimental economist interested primarily in the foundations and the methodology of social

science. He teaches in the Department of Philosophy of the University of Milan (Italy). A believer in interdisciplinary research, he makes an effort to publish both in philosophical and scientific journals. His first book, *The methodology of experimental economics* was published by Cambridge University Press in 2005. In 2011 he co-edited with Daniel Steel *The philosophy of social science reader*, which was published by Routledge. His most recent monograph is *Understanding institutions: the science and philosophy of living together*, which was published in 2016 by Princeton University Press.

Contact e-mail: <francesco.guala@unimi.it>