Erkenntnis (2023) 88:2673–2694
https://doi.org/10.1007/s10670-021-00472-z

ORIGINAL RESEARCH

Norms^2: Norms About Norms

Chiara Lisciandra[1]

Received: 24 June 2020 / Accepted: 25 September 2021 / Published online: 9 November 2021

Abstract
In this paper, I outline and defend the view that variations in compliance levels with one and the same norm represent different *norms about following norms*. In support of this claim, I first argue that classic game-theoretic accounts, which define norms as Nash equilibria of noncooperative games, typically consider variations in compliance levels as separate norms. After that, I suggest a more fine-grained, game-theoretic distinction that accounts for degrees of compliance with the same norm and I show how to incorporate such an account into a psychological framework. Finally, the paper examines what given degrees of compliance can reveal about the dynamics underlying the process of norm change. I will argue that they are indicators of different reactions to the introduction of new norms.

1 Introduction

This paper focuses on a specific aspect of norms, which concerns the tendency that different groups exhibit towards norm compliance. By this, I mean the meta-norms that regulate norm compliance, or what I call *norms about following norms*. These norms regulate the way in which the prescriptions of a norm apply to a certain situation. Examples are ubiquitous: think for instance of norms like traffic rules or punctuality. Some groups tend to comply strictly with such norms, whereas others tend to be more lax. What varies across these groups is the strength with which the prescriptions of the same norm apply to a certain situation.[1]

Even though it is common to characterize groups according to their attitudes towards norms, e.g., people from Switzerland are typically punctual, people from

[1] Note that the tendency of groups to adhere to norms does not have to be the same across all instances of norms, but may depend on the sphere of behaviour that the norm regulates: while certain groups may be particularly strict with, e.g., kinship norms, others may be more strict with civic norms. More on this below in Sect. 5.

✉ Chiara Lisciandra
c.lisciandra@rug.nl
http://www.chiaralisciandra.com

[1] Faculty of Economics and Business, University of Groningen, PO Box 800, 9700 AV Groningen, The Netherlands

Italy tend to drive recklessly, the idea that there is a kind of group "propensity" to follow norms is contentious in the literature. A different view is that groups have different norms entirely, e.g. certain groups have a norm that requires them to arrive exactly on time, while others have a norm that allows them to arrive some minutes after the agreed time.

The difference between the two positions can be illustrated by means of an example. An Italian saying goes like this: "In Milan traffic lights are instructions; in Rome, they are suggestions; in Naples, they are decorations." Even though this statement is certainly an exaggeration, is it the case that people in Milan, Rome and Naples have one and the same norm, but the norm is taken more or less seriously in different places? Or do people from the three cities have different norms entirely–a red traffic light represents different norms in the three cities, such as "never cross", "cross sometimes", "cross if you like"? And why does this matter?

The two views just mentioned:–one, that different groups have one and the same traffic norm but different compliance levels, and the other, that they have different traffic norms altogether–reflect two competing positions in the literature. On the one hand, it is typical in cross-cultural psychology and institutional economics to provide scales that rank groups with respect to, e.g., their tendency to violate norms or to abide by the rules of society (see, e.g., Gelfand, 2011). On the other hand, formal treatments of strategic behaviour, and in particular the tradition that goes back to the theory of games (see, e.g., Lewis, 1969), typically subscribe to the view that people from different groups have different (traffic) norms altogether. The incentives of the strategic games are different in Milan, Rome and Naples and they are reflected in the behaviour of the group members.

The picture that emerges from the game-theoretic approach seems to be at odds with the picture from cross-cultural psychology and institutional economics. If different groups have different norms entirely–in this case different traffic rules–then saying that some groups abide by the rules more strictly than others is unwarranted–each group is simply following the norms of their group.

In this paper, I will first show one way in which we can use resources from game theory to explain the case where groups exhibit different degrees of compliance with one and the same norm (Sects. 2 and 3). Of course there are situations where groups have different norms altogether. But, I will argue, game theory can shed light on the case where groups have one and the same norm but follow it to a different extent. In doing so, this paper will show how to reconcile different research programs–game theory on the one hand and cross-cultural psychology/institutional economics on the other–which prima facie seemed to be conflicting.

Unifying theories that were so far apart has its own philosophical merit; but–it might be argued–this is mainly a theoretical enterprise. After all, we might still be left wondering whether, when we compare the behaviour of the members of different groups, we can infer whether they have different traffic norms entirely, or one and the same traffic norm but different compliance levels.[2]

[2] The question whether, when we observe someone seemingly violating a norm, we can say whether he/she is free-riding or following a different norm is an instance of Wittgenstein's rule-following paradox. According to the paradox, "no course of action could be determined by a rule because every course of action can be made out to accord with the rule" (1953, #201). In this paper, I am following the literature

To answer this question, I will first show that an analysis of the psychological motivations underlying norm compliance provides a testbed for the aforementioned distinction. To this purpose, I will draw on Bicchieri's theory of norm compliance (2006) and show that it contributes resources to disentangle the case where different groups have different (traffic) norms altogether, from the case where they simply have one (traffic) norm but different compliance levels (Sect. 4).

After that, I will show that such a distinction can be analysed in the context of norm change–the process that groups undergo when a change is made to the norms they live by (Sect. 5). I will use evidence from empirical work on norm change to show that different groups do indeed have different *norms about norms*. To see this, consider once again the example of Milan, Rome and Naples. Suppose that we could observe the reaction of the members of the three cities to the introduction of a new norm. If they had different norms entirely, and the behaviour of their members reflected the respective incentives, we would expect them to respond to the new incentives in a similar manner. In other words, we would expect the three groups to move to the new norm as new incentives apply.

On the other hand, if they have a different tendency towards norm compliance, if they have different *norms about norms*, we would expect them to respond to the new norm in different ways: more specifically, in a way that reflects their overall tendency towards norm compliance. The working hypothesis in this paper is that groups that have a particularly strong norm about norm-following, also exhibit a sort of allegiance towards existing norms, which makes them more resistant to replace their current norms. To show this, in the final section of this paper I will discuss a case study in institutional economics from a paper by Fisman and Miguel (2007) concerning a change in traffic regulations affecting diplomats working at the United Nations in New York.

Overall, the kernel of the argument defended in this paper is that *norms about norms* indicate the tendency that groups exhibit towards norm compliance, i.e. the seriousness with which they take the prescriptions of the norms. By showing that the notion of norms about norms is game-theoretically grounded, this paper brings together different approaches to the study of norm compliance and suggests how to combine them in a unified picture. Further, the paper presents a way in which to test the hypothesis that norm compliance comes in degrees, that these degrees represent the norm about norms of a group, and that such degrees clarify the interpretation of facts about norm change.

Footnote 2 (continued)

that proposes a solution to the paradox by appealing to the role that the community plays in interpreting behaviour (see e.g., Sillari, 2013): when the majority of the members of a group "violate" a norm to a certain extent, this indicates what is *normal* for the group to do, or in other words, what is the group's norm about norm-following. Bicchieri's introspective analysis offers yet another way to investigate this issue, and in this paper, I apply her framework specifically for this purpose.

2 State of the Art

The literature on norms is vast and cuts across different disciplines—from philosophy to economics, law, psychology and anthropology–just to mention the main fields. While this variety of approaches is fruitful in clarifying different aspects of norm compliance, at the same time, the literature is divided on the issue that is central to this paper, which is, how to account both theoretically and empirically for differences in norm compliance across groups.

At one end of the spectrum, cross-cultural psychology and institutional economics provide measurements of degrees of compliance, which highlight the differences that different groups and cultures exhibit in their norm-driven behaviour (see, e.g., Andrighetto et al., 2016). The empirical literature in this field has focused on the development of indicators that compare aspects such as the level of norm abidance, corruption or trust across countries. Examples include the World Bank's "Global Governance" indicator, which encompasses dimensions such as "Control of Corruption" and "Rule of Law", or the "Prevalence of Rule Violations" index (Gächter and Schulz, 2016).[3] These studies implicitly endorse the view that norm-compliance comes in degrees.

On the other end of the spectrum, classic game-theoretic accounts, which define norms as Nash equilibria of noncooperative games, typically consider variations in compliance as separate norms. More specifically, within the standard game-theoretic framework, norms are defined as solutions to interaction problems or, more formally, norms are Nash equilibria of repeated noncooperative games.[4]

As I will show in more detail in the rest of the paper, this specific game-theoretic framework does not enable cross-group comparison of compliance levels with one and the same norm. Obviously, the theory admits variations in players' behaviour, depending for instance on the individual level of internalisation of a norm, or on mechanisms such as the "trembling hand" or randomisation. However, the norm that results from repeated noncooperative games, in order to be a norm, has to be an equilibrium of the game. Within this framework, an equilibrium incorporates a combination of compliance, violations and sanctions. Because violations are constitutive of a norm, i.e. they are part of its definition, different degrees of violation represent different norms altogether. Therefore, information about violations cannot provide

[3] Another well-known example in the literature is the work by Gelfand (2011), which makes a distinction between "tight" and "loose" societies. This distinction measures another aspect than "norms about norms". By "tight" and "loose" societies, Gelfand and her collaborators mean that some groups tend to crowd the space of social interaction with more norms than others. In a way, their scale reflects the amount of norms there are in a group. Accordingly, "loose" societies may look more permissive than others because they have fewer norms; but they may still apply their norms very strictly.

[4] Clearly, game theory is a set of analytical tools, not just a single theory. It includes—among others—classical, evolutionary, epistemic game theory, cooperative and noncooperative game theory, and games with complete or incomplete information. Further, Nash equilibrium is just one solution concept alongside other concepts such as correlated equilibria and subgame perfect equilibria. In what follows, I will refer to classical game theory to indicate that part of the theory that defines norms as Nash equilibria of noncooperative games (on this characterisation, see Gintis, 2010 and the literature therein).

at the same time an independent standard to measure levels of compliance across groups.

At first sight, the picture that emerges from the game-theoretic approach seems to be at odds with the work on cross-cultural psychology and institutional economics mentioned above. In the next section, I will show that it is possible to reconcile these approaches by adopting a specific account in game theory, which is known as the "rules-in-equilibrium" account. This account has been developed by Guala and Hindriks (2014) and Hindriks and Guala (2015). It draws on Gintis' work on social norms as "choreography" (2010), which in turn derives from Aumann's concept of correlated equilibrium (1987). I will argue that the rules-in-equilibrium account lends itself to drawing a distinction between different levels of compliance, because it equips us with an independent reference tool that can be used for cross-groups comparison.

3 Degrees of Compliance

The purpose of this section is to show that the "rules-in-equilibrium" account provides a game-theoretic distinction that accounts for degrees of compliance with one and the same norm. I will first present the theoretical roots of the "rules-in-equilibrium" account, which are the "rule-based" account and the "equilibrium-based" account from game theory. I will argue that the equilibrium-based account treats differences in compliance across groups as separate norms altogether, whereas the "rules-in-equilibrium" account allows us to treat such differences as degrees of compliance with one and the same norm.

First, some terminological distinctions that will be used throughout the paper. While a variety of definitions of norm have been suggested in the literature, I shall use the term norm in its broadest sense, to refer to the prescriptions that regulate the interactions between the members of a group. Examples of norms are laws, conventions, etiquettes, and social norms. Norms are shared by (a sufficient number of the) members of a group and have a normative component, which indicates what should be done in an interactive situation, either because it is considered to be the right thing to do or because it is expected by others and there may be sanctions (see on this, e.g., Bicchieri, 2006; Brennan et al., 2013).[5]

Norms include both formal and informal agreements. In the former case, they are statements that are typically written in a body of laws or in codes of conduct (Hart, 1961). In the latter case, they are informally transmitted between the members of a

[5] In the literature, there has been an extensive discussion of different taxonomies of norms. Typically, conventions are distinguished from social norms because the former prescribe behaviour that is convenient for individuals to follow; by contrast, social norms require the giving up of an individual gain in favour of a group gain. However, some authors have challenged this distinction, by showing that conventions have a normative component as well. Although differences of opinion still exist, I subscribe to the latter view, which has been defended in Guala (2013), that conventions have a normative force that motivates action, even when it would be convenient to deviate from their prescriptions.

group. Social norms such as greetings or queuing, are typical examples of informal norms.

In this paper, norms about norms are defined as the meta-norms that prescribe how strictly the prescriptions of norms should be followed. Conceptually, there may be just one meta-norm that applies to the entire set of norms or there may be different meta-norms for specific categories of norms. Whether the former or the latter holds is an empirical question that can be assessed on a case-by-case basis.

With this terminological clarification in place, I will now briefly summarise the contents of the rule-based account and the equilibrium-based account and discuss how they relate to each other. After that, I will suggest how both accounts can be modified to accommodate the notion of norms about norms.

3.1 The Rule-Based Account and the Equilibrium-Based Account

A paradigmatic example of the rule-based account comes from the work of the institutional economist Douglass North (1990, 1991). North developed his account with reference to institutions, but his approach can be generalised more broadly to other instances of regulated social interactions. The main idea underlying his approach is that in order to solve coordination or cooperation problems, individuals have introduced *rules* that facilitate interactions between the members of a group. By imposing costs for deviation from such rules, individuals have at the same time an incentive to stick to their prescriptions. In his own words: "Institutional constraints include both what individuals are prohibited from doing and, sometimes, under what conditions some individuals are permitted to undertake certain activities. [...] [They] [i.e., institutional constraints] therefore are the framework within which human interaction takes place." (1990, p. 4)

This account adequately captures the way in which rules signal the "appropriate" behaviour in a specific circumstance. This is crucial because there are often several different ways in which we can solve interaction problems, e.g. driving on the left or the right, handshaking or bowing, using coins or shells as currency. According to North, rules dictate which of the possible alternatives applies in any particular context and thus reduce the uncertainty that would otherwise characterise our social interactions.

The account, however, has been criticised precisely because it seems to overstate the role it attributes to *rules*: it is often not sufficient (nor it is necessary) for a rule to be stated, either formally or informally, for it to be followed. In other words, there are several situations where a rule is known by the members of a group, even though it is not complied with.

An account that better accommodates this aspect is known as the equilibrium-based account. It has an important tradition in economics, e.g., Schotter (1981), and in philosophy, e.g., Lewis (1969). According to this account, social interactions can be represented as noncooperative game-theoretic structures whose Nash equilibria correspond to the norms of a social group. The choices in equilibrium are the best response to the actions of other individuals. Thus, violations, i.e. behaviour out of equilibrium, are unstable and will most likely disappear. Like North, Schotter also

refers to institutions, but Schotter's focus is more specific in that it only includes noncooperative games, i.e. games that are not regulated via communication or contracting agreements.

One of the problems with the equilibrium-based account is that it, too, fails to recognise one of the distinguishing features of norms: there are in fact several interactive games whose Nash equilibria may be selected, even though this does not imply that a society has solved an interaction problem by means of norms, or an institution. Defection in the prisoner's dilemma, to mention just one of the most extreme cases, is an equilibrium that does not represent an institution, but rather the lack of one.

3.2 The Rules-in-Equilibrium Account

Guala and Hindriks (2014) have shown that it is possible to combine the features of the two previous accounts in a synthesis that preserves the best aspects of both approaches. This is known as the rules-in-equilibrium account.[6] The way it works is that it considers the solutions to interaction problems as equilibria that are prescribed by certain specific rules. In this sense, rules act as equilibrium selection devices: they correlate people's behaviour and guide it towards the choice of an option that is convenient for everyone to follow.

The idea of correlated equilibria goes back to Aumann (1987), who presented it as an extension to Nash equilibria of strategic games: the discussion is rather technical, but the basic idea is that the participants in a game select a conditional strategy, namely a strategy that prescribes different actions on the basis of a random variable.[7] For instance, the outcome of a coin toss tells players to select a particular action in a game with multiple Nash equilibria, such as action S_1 if the toss gives heads, action S_2 otherwise. Once an equilibrium is in place, individuals tend to stick to the related strategy and would not want to deviate from it, because the equilibrium offers a solution to an interactive game. If we now transfer the framework of correlated equilibria to the context of the norms of society, where individuals face coordination problems continually, the idea is that a coin toss provides the "first" signal on how to play the game. In repeated interactions, by a process of gradual adjustment, a certain equilibrium emerges that also includes violations and sanctions (more on this below).

Having outlined the basic features of these different accounts, we can now consider the question that is central to this paper, which is how to compare cross-group differences in levels of compliance with one and the same norm. To do so, I will first argue that the equilibrium-based account does not allow for cross-group comparisons and, then, that the rules-in-equilibrium account does so.

According to the equilibrium-based account, norms are equilibria of *repeated games*. Repeated games describe norms as equilibria that consist of a specific

[6] By rules, the authors mean signlas, i.e. equilibrium selection devices.

[7] For a more technical discussion see Gintis (2009, 2010) and Vanderschraaf (1995, 1998, 2001).

combination of compliance, violations and sanctions. The role of violations and sanctions is crucial for the definition of a norm; in other words, the extent to which a norm is violated and the extent to which violations are sanctioned are constitutive of the definition of a norm. A formal result from game theory, known as the *folk theorem*, shows that there can be an infinite number of combinations of compliance, violations and sanctions, all of which form an equilibrium strategy. This depends on the fact that norm compliance presupposes an effective system of sanctions, but sanctions can be established in an unlimited number of ways: the minimum requirement is that, in the long run, the system succeeds in making the benefits of compliance greater than the benefits of mutual defection. *Tit-for-tat*, for instance, is a strategy of a repeated game, where each violation is punished by a sanction that restores compliance (Fudenberg and Maskin, 1986).

In this way, a game-theoretic account explains different equilibria as different norms. There may be a group where, for instance, every deviation is systematically followed by a sanction. Then *that* is the equilibrium of the game, i.e. the norm. There may be another group where only some deviations are punished; then *that* is another equilibrium of the repeated game, i.e. a different norm. The two groups have two altogether different norms; thus, strictly speaking, we cannot compare them with each other.

By contrast, the rules-in-equilibrium account has an advantage over the equilibrium-based account, or so I argue. The rules-in-equilibrium account allows us to compare groups with respect to their degree of compliance, by providing information about the equilibrium selection device and the related rule. The idea behind correlated equilibria is that signalling devices provide an anchoring mechanism that allows individuals to select one option among several possible alternatives. One example of a signalling device is the history of the game that the members of a group play (Lewis, 1969). Another example is traffic lights. They prescribe the way in which drivers should behave at an intersection and exclude in this way other possible coordinating systems. As we know, however, in some countries their prescriptions are followed strictly, in other countries more loosely. This is because rules steer individual behaviour and indicate the way in which the members of a group should behave in a certain situation, even though they do not entirely determine behaviour.

To see that this is the case, notice that if rules were to correspond exactly to the equilibrium strategy, then the definition of a rule should incorporate the expected degree of compliance. For instance, there would be a rule that it is permitted to cross the street with a red traffic light, e.g., two thirds of the times. However, this is not what rules do: rules simply state what should be done, not the frequency with which we can deviate from their prescriptions.

Notice also that if one and the same signal, e.g. a traffic light, is compatible with different equilibria, we can measure differences in compliance levels across groups. This is because the rule prescribed by the signal provides the benchmark against which compliance, violations and punishment can be measured. If we stick to the standard equilibrium-based account that considers norms merely as equilibria of compliance, violations and sanctions, then behaviour out of equilibrium is behaviour that does not conform with that pattern; and since violations are part of the equilibrium strategy, they are not strictly speaking *deviations from the norm*. Nevertheless,

they are deviations from the rules, i.e.,violations of the prescription of the signal. This distinction, which we can make in the rules-in-equilibrium account, is crucial as it provides information about levels of compliance. In other words, this is the key to comparing groups with respect to their tendency to follow norms. And by distinguishing between the signal and the equilibrium, we can preserve information both on what the the signal prescribes and on how different groups follow its prescriptions.

The reason why the distinction is important is that when we talk about groups or countries where corruption is high, or where institutions are fragile, we are talking about groups that are violating prescriptions that say, for instance, that citizens should pay taxes or that they should not bribe government officials. If we treat those situations as if they were instances of norms that allow citizens regularly to evade taxes, or repeatedly to bribe officials, then it becomes difficult to talk about transgression at all, or to compare groups in terms of compliance.[8] Not only is this a relevant conceptual distinction but also, as I will show in the next two sections, it has consequences for the ways in which groups change the norms they live by.

4 Norms About Norms

The previous section identifies the rules-in-equilibrium account as a theoretical framework that enables us to classify groups with respect to their degree of norm compliance. As was briefly mentioned in the conclusion of the section, if we endorse this view, we can compare groups with respect, for instance, to their level of corruption; this would not be possible in the equilibrium-based account. According to that analysis, those groups where corruption is widespread would have a norm that permits them to be more corrupt than others. In that sense, citizens of corrupt societies would be very strict norm followers; they would just be following a norm that allowed them to be corrupt most of the time.

That interpretation, however, clearly conflicts with the way in which we commonly intend the meaning of compliance, corruption or even honesty. To stay with the example of corruption, in the equilibrium-based account it does not even make sense to talk about corruption as we usually do. There is no external basis of comparison according to which some groups are more corrupt than others, since each of them would be violating a norm from the perspective of the other group. Note that it would be self-contradictory for an advocate of the equilibrium-based account to say that corrupt countries are those where citizens violate laws more often than others, because laws without a corresponding equilibrium are silent, they are just empty statements. Even though it is common parlance to say that corruption occurs when

[8] To stick with the example of tax evasion, if upon comparing countries with each other, we observe that in one country tax evasion is significantly less frequent than in another country, then we can conclude that the members of the first country have a stronger norm against evading taxes than members of the second country. In other words, when the behaviour of a group is consistently less compliant with the norm than the behaviour of another group we can explain it by means of different norms about norm compliance.

corrupt actions are the equilibrium, strictly speaking, there is no benchmark within the domain of the theory that can justify that statement.

By contrast, within the rules-in-equilibrium account, the rule on the basis of which an equilibrium emerges "tells" the members of a group what they are allowed to do in a certain situation. This is relevant for an assessment of how "seriously" the prescription is taken. To see this, suppose that we were in a country where it is common to cross the street even if the traffic light is red. Compare this situation with a country where this happens only rarely. In both cases, it is the distance between the rule and the equilibrium that provides a measure of the extent to which individuals will stick to the rule as against the opposite.

So far, the question of how to account for degrees of compliance with one and the same norm has been approached mainly theoretically. However, this notion can be made psychologically sound by identifying cases where the members of different groups recognise that the same norm is in place, but treat it with different degrees of "seriousness". In other words, if groups that tend to be described as more corrupt than others, in fact perceive their behaviour as transgressing, even though that is usually tolerated, then this would be a way to show that the same norm can be described as having different compliance levels.

To discuss this point, it is helpful to move from a purely game-theoretic account to a psychological account. In this area, one of the approaches that have crucially contributed to our understanding of norm-abiding behaviour is Bicchieri's theory of norm compliance (2006). In what follows, I will focus on the psychological side of Bicchieri's theory to show that, within that framework, we can show that there are cases where we can talk of different compliance levels.

Bicchieri's account (2006) refers in particular to a specific group of norms, i.e. social norms, but we can follow her in thinking about norm compliance more generally. Bicchieri's account allows us to assess whether a social norm N exists in a group, by eliciting individual beliefs and expectations: if the individuals of a group recognise that a social norm applies to a certain circumstance, and they believe that (a significant number of) others will follow it and that (a significant number of) others expect him/her to comply, then we can conclude that that norm is in place.

Bicchieri's theory refers to a *significant* number of group members, thus admitting variation to full compliance. Clearly, there may be groups whose members believe and expect that almost the entire population will follow the norm, and groups whose members believe and expect that only a fraction of the population will follow the norm. But as long as the fraction is above the threshold of significance, Bicchieri's theory considers the related behaviour to be a social norm. Different fractions of compliance represent different degrees of compliance, i.e. norms about norms. For each fraction, the beliefs and expectations of the individuals are accordingly tuned: the members of a group know that there is a norm, yet it is not fully complied with, and so a certain level of deviation is admitted.

Let us see how Bicchieri's account works in practice. As an example of an informal rule, consider the case of punctuality. There may be groups whose members strictly stick to the rule and others where some members tend to arrive a bit later than the agreed time. Arguably, the latter do not have a different rule than the former; they may be simply more self-indulgent in their behaviour.

To assess whether this is the case, we can elicit beliefs and expectations to attest that a norm of punctuality is in place, irrespective of whether it is strictly followed. All that is needed for a norm to be in place is that the members of a group recognise that by not being punctual, they are disregarding—even only slightly—the expectations of their fellow group members. Within Bicchieri's account, the only situation where a norm is not in place is when individuals are not even aware that the norm applies, they do not have related beliefs and expectations and—at least for certain individuals—there are no related sanctions.

The same kind of analysis can be applied to formal rules. Think, for instance, of the traffic rules for crossing the street at a traffic light. There are cities in the world where people indulge in jaywalking, but this does not mean that citizens are not aware of traffic rules. Bicchieri's test gives us a way to elicit beliefs and expectations in order to observe whether citizens are aware of the rules and of their own non-compliant behaviour. This test can be used precisely to distinguish between the case where i) an individual violates the traffic norm for personal reasons, for instance because he/she wants to appear confident in the eyes of others, or challenge authority; and ii) the case where an individual is aware that by her behavior she is breaking a traffic rule but at the same time thinks that the violation is not serious, as that level of deviation is usually tolerated in her group. The latter scenario is crucial as it identifies what is the norm about norm compliance in a certain group.

The empirical literature on social norm compliance already shows that there are individual differences in sensitivity to a norm, i.e. in the strength with which a norm applies to a given circumstance. According to Bicchieri, norm compliance is indeed subject to individual differences: in her model, a parameter k expresses the sensitivity of an individual to norm compliance. A similar idea was proposed by Crawford and Ostrom (1995), where the authors introduce a parameter to express the deontic power of prescriptions, i.e. the perceived rewards and costs of obeying a norm or breaking it. Guala (2016) also refers to the role of deontic power when discussing the normativity of institutions.

Rather than focusing on sensitivity to the rules at the individual level, the account I propose considers that the tendency to comply with norms is a group characteristic, which mediates individual differences. Certainly, there are differences in sensitivity among the individuals of the same group; but since norms are social phenomena, it is plausible to expect that in-group characteristics are more homogeneous than out-group characteristics. And such in-group features represent what I have called *norms about norms*.

Ultimately, this tendency can be explained in terms of individual characteristics; it depends on how seriously each individual takes the prescription of the norm. However, since becoming part of a group where, e.g., there are more "norm-benders" increases the likelihood that new members will bend the norms as well, it is possible to see that this feature is a group/characteristic, i.e. that there is a norm that affects the way in which individuals, *qua* group members, behave with respect to norm compliance.[9]

[9] To see this, take for instance a conformist, i.e. someone who follows the norm of his/her group almost all the time. Arguably, the prevalence of conformists in a group cannot just be accounted for in terms of individual psychological traits, i.e. this would imply that there are groups where unexplainably the

Beyond Bicchieri's test, there is yet another way to assess the relevance of differences in compliance, which is to show that they are a good predictor of norm change. The aim of the next section is to show that groups that have different norms about norms react differently when exposed to new norms or to policies directed at behavioural change. In other words, I will argue that different norms about norms characterise the ways in which groups go through processes of norm change.

5 Case Study: An Empirical Analysis of Norm Change

Suppose that we had a scale that measures degrees of compliance across groups. Let us call this scale the $(norms)^2$-scale. Let us also say that the higher the value on the scale, the closer the alignment between the prescription of the norm and its application in practice. According to this scale, some groups are strict "rule-followers", others are "rule-benders", depending on their different *norms about norms*.

The aim of this section is to consider in more detail the relation between degrees of compliance and reactions to norm change. The main questions for analysis are: How does a strong degree of compliance combine with norm change? Is one feature detrimental to the other? Or are those groups, who tend to follow norms more strictly, also quicker at replacing their own norms?

Answering these questions is crucial because if degree of compliance predicts an inverse relation to norm change, i.e. if those groups who tend to follow norms more strictly, are also slower at abandoning their own norms, then we have a major theoretical reason for positing the existence of *norms about norms*. This is because, having a strong norm about norms can explain why the members of a group are also more strongly "attached" to their own norms: they have a sort of inertia towards existing norms. By contrast, if groups with a high degree of compliance do not show any particular resistance towards norm change, and switch swiftly to a new norm, then we would not have major reasons to posit a norm about norm compliance. We could explain that the members of a group follow new norms in a similar way to that in which they follow existing norms, for instance because of the incentives they have to do so.

Testing the hypothesis above is important not only for theoretical reasons but also, more practically, because there are situations in which it may be beneficial for a group to change the norms they live by. For that purpose, it may be helpful to explore ways of moving a group towards a different norm; and to increase awareness of the dynamics underlying norm change.

Before addressing the questions stated above, two specifications are needed. First, note that it is unlikely that an overarching meta-norm exists that covers the entire set of norms. A group may have more than just *one* norm about norms, depending on the kinds of norm at stake. A norm of compliance may apply to clusters of norms

Footnote 9 (continued)
majority of people have the same character trait, but rather in terms of being part of a community with a particular norm about norm-following.

that are similar in kind, as for instance norms that regulate behaviour in the public sphere, or norms that regulate kinship behaviour; it is also plausible that those groups where allegiance to public sphere norms is high will have low allegiance to kinship norms and vice versa (on this point, see Knack and Keefer, 1997).

Secondly, the idea of norm compliance as a group-level feature draws on the literature on comparative research which aims at classifying cultures and societies on the basis of traits, such as values, beliefs, cultural attitudes, and institutional quality. In order to measure such traits, researchers draw on survey studies that typically involve thousands of respondents from different representative groups across several years. On the basis of the data they collect, researchers build aggregate indicators that capture the prevalence of a particular trait at the group level.[10] For instance, Hofstede's individualism/collectivism dimension classifies societies according to their tendency to endorse individualistic or collectivistic values. In relation to these values, members of individualistic societies tend to endorse norms that lead to the pursuit of individual interests, whereas collectivist societies tend to rely more on family ties and place the relevant group's interest over individual interests.

Clearly, aggregating data at a high level of generality implies excluding individual differences that may be relevant, especially depending on the kind of analysis undertaken. In the context of this paper, the main point is that the literature offers a variety of sets of measurements which classify groups according to their tendency to endorse values that are associated with, and expressed by, sets of norms, for instance kinship norms or civic norms.

Now, recall the example from the first section about Milan, Rome and Naples and the question whether the residents of the three cities have each a different propensity to follow the same set of norms. If that was the case, we would expect that when an old norm changes, the way in which a new norm is adopted will reflect the general attitude towards norms. More precisely, how would this attitude affect the response to norm change?

The main hypothesis of this section is that those groups who strictly comply with norms tend to stick to their norms more firmly than others. In other words, the members of such groups are so compliant with their own norms that they show some reluctance to deviate from their prescriptions. Similarly, groups of norm benders might also show some resistance to comply with new norms, as they have a weak norm about following norms. By contrast, those groups who are in between strict norm-followers and norm-benders may respond more quickly to the introduction of new norms or to norm change more generally. In other words, they may be more

[10] An important question in this respect is what aggregation method works best to represent the norms of norm-compliance. There may be groups where a norm is very rarely violated, but when it is violated, the violation is severe. And there may be groups where a norm is more often violated, but less severely. The question arises, which measure captures more closely norm-compliance as a group-level characteristic. This paper endorses the idea that the frequency of instances of violations is a better indicator of norms about norm-compliance, but alternative measures are worth exploring together with the robustness of their outcomes.

flexible about changing norms, even though in the long run they may not exhibit high levels of compliance or deviation.[11]

The remainder of this section will explore the hypothesis above by focusing in particular on differences in compliance levels with civic norms, but the analysis might be extended to other kinds of norms. Examples of civic norms are e.g., norms against cheating on taxes, littering, avoiding a fare on public transport, claiming government benefits to which the claimant is not entitled, etc. Civic norms typically correlate with low levels of corruption: they tend to facilitate economic transactions and increase citizens' cooperation and trust (Knack and Keefer, 1997). In what follows, differences in groups' adherence to civic norms will be compared with their responses to the introduction of a new civic norm.

A study whose data can be used to observe the relation between adherence to civic norms and norm change is a paper by Fisman and Miguel (2007), where the authors investigate a correlation between the tendency to violate norms in a foreign country and corruption in the country of origin.

The authors analyse data on the parking fines incurred by international diplomats from 149 countries working at the United Nations in New York between 1997 and 2005. The data on corruption is taken from an indicator built by Kaufmann et al. (2006), the World Governance Indicator (WGI), which is a compound measure of the quality of governance from 1996 to the present day. One of the dimensions of the indicator is the corruption index, which measures individual perceptions of the abuse of power and efficiency of the state.

In what follows, the corruption index is used as a proxy for the $(norms)^2$-scale, as it indirectly captures the cross-country degree of compliance with some of the norms that belong to the civic sphere. Even though the index does not directly measure adherence to civic norms, it features as a valid, if imprecise, substitute for a *norm about civic norms* because typically low levels of corruption are also expressed by citizens' compliance with civic norms.[12]

Until the end of 2002, diplomats and their families in New York enjoyed an immunity status that enabled them to avoid paying parking fines. Immunity does

[11] In a recent work on norm change, Bicchieri (2017) argues that one way in which the process of norm change can get started is via trendsetters, i.e. individuals with low sensitivity to the norms, who–precisely because of this feature–are particularly prone to break existing norms and help establish new norms. Unlike Bicchieri, in this paper I am focusing on group characteristics rather than on individual personal traits. I consider how the process of norm change develops depending on group's attitude towards norm compliance. For instance: how does a new civic norm get established in groups where typically compliance with civic norms is lax? My claim is that it is unlikely that low sensitivity to norms is optimal for norm change, because individuals with low sensitivity will probably disregard both old, current and new norms. Ultimately, however, this is an empirical question and with the case study that follows, I aim to provide an example that supports the mechanism I describe (on this point, see, e.g., Paluck, 2009).

[12] The six dimensions of the WGI are Voice and Accountability, Political Stability and Absence of Violence, Government Effectiveness, Regulatory Quality, Rule of Law, Control of Corruption. The dimensions highly correlate with each other and it might hence be asked whether the corruption index indirectly measures a norm about civic norms or rather citizens' confidence in the ability of the state to enforce laws and civic norms. This point will be discussed again later on. For the moment, it suffices to say that law and norm abidance do not depend only on strict enforcement mechanisms and control; an important role is also played by what is called the normativity of the law, i.e., by the way in which citizens respond to the prescriptions of the law. It could be the case that the respect that citizens exhibit

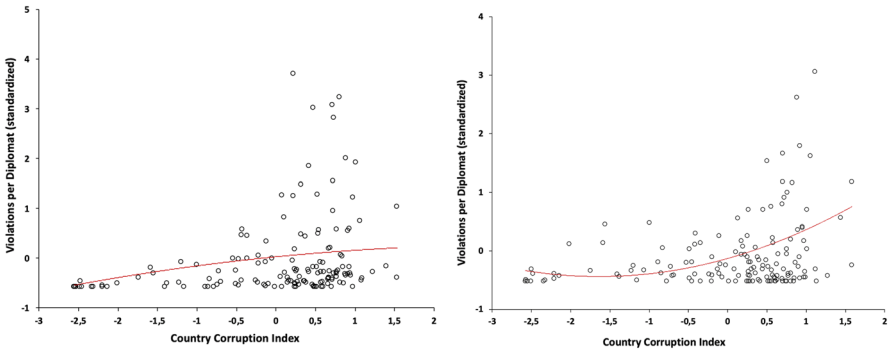

Fig. 1 Relation between country corruption measure and violations per diplomat before the enforcement of the 2002 law (on the left) and between country corruption measure and violations per diplomat after the enforcement of the 2002 law (on the right). The dots represent for each country the number of violations per diplomat

not, strictly speaking, imply that the law did not apply to diplomats. In fact, diplomats got traffic tickets anyway; they were just not obligated to pay them. When immunity was removed, what changed was the implementation of the law. This can be considered a case of norm change, as a crucial component of the norm was changed, i.e. the way in which it was enforced. First, the authors show that diplomats coming from countries that are more corrupt violated parking rules more often than diplomats from less corrupt countries. The graph on the left of Figure 1 represents the situation before 2002.

Secondly, the authors focus on what I shall call the enforcement period, which started when immunity was revoked in 2002. As the authors show, the number of violations significantly dropped, even though the main pattern remained. The graph on the right of Fig. 1 represents the situation after 2002.[13]

What is important to observe here is what happened before and after the change in the implementation of the law, as this gives us an opportunity to consider how diplomats from different countries responded to this change. Even though the data was not collected for the purpose of analysing norm change, we can use it to examine the relation between countries of origin and the tendency to change behaviour.

In both cases, the graphs show a positive correlation between the corruption index and the tendency to violate the norm. Further, the graphs show that, even though violations dropped significantly in the enforcement period, the response was not identical, i.e. while certain countries' diplomats stopped getting fines, others kept violating the norm to a certain extent.

Footnote 12 (continued)
towards laws and norms originally stems from the threat of sanctions, but it is usually acknowledged that the seriousness with which citizens regard the law equally contributes to law abidance.

[13] Note that the data have been standardized to allow for direct comparison of the two cases, because of the different magnitudes of the scales of violations before and after 2002.

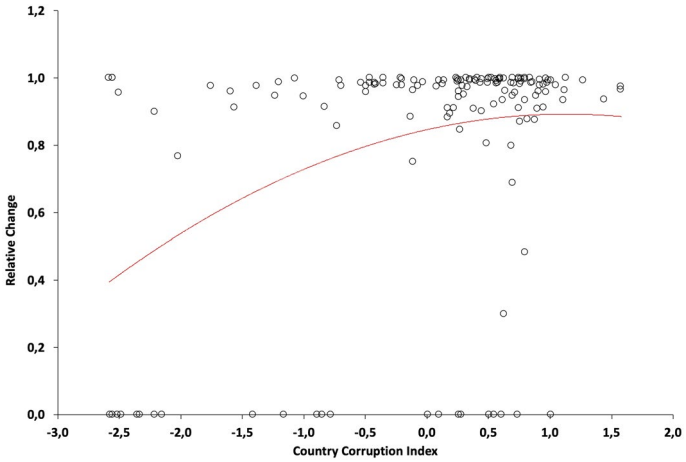

Fig. 2 Relative change between pre-enforcement and post-enforcement periods

If the intervention had had consistent effects, and the groups had had different norms entirely, we would expect that each group would refrain from violating the norm according to the new incentives. Given that the responses were not analogous, in what ways did the countries' diplomats change their behaviour? To measure the extent to which groups changed their behaviour in the enforcement period, we measure relative change, i.e. how much the behaviour changed in proportion to what it used to be before the enforcement of the law. This measurement of relative change captures the idea that if the diplomats of two countries reduce their violations by e.g., 10 units, but one started from e.g., 100 violations and the other from 20 violations, then the latter group has changed its behaviour in a more important sense, because its diplomats reduced their violations by 50%.

The graph in Fig. 2 shows relative change as a function of the corruption index.[14] The corruption index ranges from [−2.6 to 1.6], where the lower the value, the less corrupt a country is. The function has a local maximum at around 0.8 of the country corruption measure, which suggests that the biggest change in behaviour occurred not at the extremes of, but in between the maximum and minimum values for corruption. To be sure, the diplomats of countries that are strict norm followers did not change their behaviour considerably, because they were already aligned with the norm. However, the important feature to observe is that the behavioural change did not occur in the same way across countries, but is more pronounced for those in between the extremes.

The analysis gives a preliminary indication that there is no linear relation between the attitude to norm compliance and the tendency to accept norm change. Another way to put this is to say that the tendency to adhere to new norms may be mediated

[14] Using (polynomial) regression analysis, the following functional relationship appears between corruption, i.e. the independent variable, and the difference in behaviour, i.e. dependent variable:

$$\text{Difference} = -0.0362 \times \text{Corruption}^2 + 0.0817 \times \text{Corruption} + 0.154 \quad (1)$$

by the attitude towards norm compliance. This result debunks the argument, that if different groups had different norms—as the classical framework suggests—then we would expect to find a linear relation between a norm and the way in which the behaviour of the members responds to its change.

Two possible objections are considered below (see Table 1 for a summary). First, an alternative interpretation of the above analysis is that the tendency we observe can be explained by the "proximity" of the norms before and after the enforcement. In other words, it might be that diplomats import the norm they follow in their own country, so that the more distant the new norm is from the old norm, the more violations we observe in the enforcement period. If this was the case, however, we would observe a linear decrease in violations. The analysis, however, does not warrant this conclusion.

To see this, take a country with a high number of traffic violations whose behaviour did not change much, after 2002. It might be said that the reason is that diplomats were still behaving as they would do in their own country, where there is no norm regulating parking. If that was the case, however, we would also expect that countries in between the extremes would demonstrate a similar tendency as before the enforcement, when they used to violate the norm "half of the time". If they were behaving as they would at home, we would expect them to exhibit a similar pattern, rather than a more pronounced shift to the new norm.

What the study presented above instead seems to suggest is that there is in fact a particular relation between compliance level and the way in which a group responds to new norms. More specifically, what the study indicates is that the tendency of groups to comply with existing norms may affect the way they comply with new norms. In particular, the idea is that, at least in the short run, countries who have a more moderate norm about following norms might be more flexible and adjust more quickly to a new norm.

A second possible objection is that the *confidence* that citizens have in the enforcement of the new norm is the main driver of norm change. Under this scenario, we would expect that when we introduce a new norm that replaces an existing norm, the pattern of behaviour in response to the new norm reflects the confidence in the other members' behaviour. Groups of strong norm followers would soon switch to the new norms, as they trust that others will do the same. Groups of norm benders wouldn't adopt the new norm, as they do not have confidence that others will follow it. And groups of medium norm followers would follow the new norm "half of the time", as they do with the norms they live by. But again, the "norm about norms" hypothesis better accommodates the evidence from the case study, for similar reasons as those presented discussing the objections above. The hypothesis predicts that groups of strong norm followers will be more resistant to changing the norms they live by because they are "attached" to the norms they have and exhibit less readiness to respond to change. Groups of norm benders won't adopt the new norm because they lack a norm about norm compliance. Groups of medium norm compliers will pick up a new norm more quickly than the other groups, because they

Table 1 Predictions of responses to norm change under the "Norms about norms" hypothesis, as against the "Proximity" hypothesis and the "Confidence" hypothesis

	Strong norm-followers	Norm-benders	Medium norm-followers
"Proximity" Hypothesis	Group members adopt the new norm more quickly than other groups because it is more similar to their "home" norm	Group members exhibit high violation of the new norm because of the large distance between the new norm and the old norm	Group members follow the new norm "half of the time", as they do with the norms they live by
"Confidence" Hypothesis	Group members adopt the new norm more quickly than other groups because they are confident that others will do the same	Group members don't adopt the new norm because they don't have confidence that others will follow it	Group members follow the new norm "half of the time", as they do with the norms they live by
"Norm about norms" hypothesis	Group members are more resistant to changing the norm they live by because they are "attached" to their norms and show less readiness to respond to change	Group members don't adopt the new norm because they don't have a norm about norm-following	Group members adopt the new norm more quickly than other groups because they have a moderate norm about norm-following

have a moderate norm about norm-compliance, which makes them follow the norms of their community without being either too lax or too strict.[15]

Overall, my observation of the relation between compliance and change has some implications with respect to the assessment of societies on the basis of their tendency to comply with norms. More specifically, it suggests that the connection that is often made between norm-compliant societies and efficiency should be reconsidered in the light of the economic cost of transition to new norms.

In the literature, societies are often considered to have well-functioning institutions to the extent that they comply with their respective norms. Furthermore, a connection is often made between well-functioning institutions and economic growth (see, e.g., Gächter and Schulz, 2016; Guiso et al. 2016; Knack and Keefer, 1997; Mauro, 2004). By well-functioning institutions, it is usually meant that individuals accept the norms of their own society and comply with them.

In the light of the analysis presented in this paper, the classifications above seem to overlook the extent to which different groups or cultures are able to combine their own norms with those of other groups. It might be the case that, beyond the allegiance to one's own system of norms, a further aspect to consider is how well societies respond to external "shocks", such as different systems of norms, for instance those introduced by foreigners or different groups. This may have consequences both in terms of economic performance and political stability. While a strong loyalty to "internal" norms may guarantee a high level of internal success, accepting different norms or adapting to new ones can also benefit a society.

The $(norms)^2$-scale might provide precise indications about what to expect from the combination of different systems of norms. It would clarify the observation that groups with different norms are likely to have clashes, by providing a measurement tool that makes it possible to predict which combinations of groups may result in more serious clashes as compared with others.

This analysis clearly does not attempt to suggest that, in the long run, a strong attitude to norms may not be a positive aspect for diverse societies, even though the process of convergence on the same set of norms will take longer. The main point is that it is likely that groups at different points on a scale that measures their *norms about norms* may have serious disagreements. This is an aspect that should be taken into account when ranking societies according to the strength of their institutions.

[15] The process by which norms change can occur gradually or abruptly depending on a number of factors (see on this, Centola et al., 2005; Kuran, 1995). For instance, if the members of a group think that the behaviour that a new norm licences is not appropriate, it will likely take longer to move to the new norm than in those cases where they believe that a norm is outdated. It would be interesting to explore whether, when keeping these factors fixed, norms about norms still mediate the process of norm change. The hypothesis is that it would anyway take longer for an individual of a group of "strict norm-followers" to switch to a new norm than an individual of a group of "medium norm followers", even when they both support, or do not support, the old norm to a similar extent.

6 Conclusions

In this paper, I have argued in favour of the claim that different degrees of norm compliance do not necessarily amount to different norms. In the first part of the paper, I have shown that the rules-in-equilibrium account provides a theoretical framework that can be used to assess the level of norm compliance within a group. In a nutshell, this account adopts the standard game-theoretic framework but attaches to it the role that a selection device plays in selecting the equilibrium of the game. I have claimed that, in doing so, this account enables us to draw a more fine-grained distinction, within which to distinguish different levels of compliance as variations with respect to one and the same norm, rather than as different and separate norms.

The theoretical analysis of this work thus contributes to filling a gap in the norm-compliance literature: on the one hand, game-theoretic accounts typically describe norms as equilibria of repeated games; and in this way, different levels of norm compliance are considered to be distinct norms. Cross-cultural psychology and institutional economics, on the other hand, build scales that rank groups according to their tendency to comply with norms; but they lack the underlying conceptual foundation that game theory provides. This paper argues that the rules-in-equilibrium account provides a unifying theory for the different approaches used in the literature. In doing so, the analysis in the current work provides further evidence in favour of the rules-in-equilibrium account and of its explanatory power.

In the second part of the paper, I have shown how Bicchieri's theory of norm compliance provides a psychological framework on the basis of which we can test the aforementioned distinction. Finally, I have shown that *norms about norms* may provide relevant information on how groups respond to norm change. First, this relation supports the idea that the distinction in question is explanatorily significant. Secondly, the results of the empirical analysis call into question the connection often made between norm-abiding societies and well-functioning societies. Whereas groups are typically classified on scales that measure compliance with existing norms, this paper argues that other indicators should be taken into account that also capture the readiness of a society to adapt to different norms or to change its own norms. One of the questions that the current literature leaves open is what it may take to change any particular *norm about norms*, so that groups may change attitudes towards norm compliance; this is a topic for further inquiry, for which this paper has hopefully laid the groundwork.

References

Andrighetto, G., Zhang, N., Ottone, S., Ponzano, F., D'Attoma, J., & Steinmo, S. (2016). Are some countries more honest than others? Evidence from a tax compliance experiment in Sweden and Italy. *Frontiers in Psychology, 7,* 472.

Aumann, R. J. (1987). Correlated equilibrium as an expression of Bayesian rationality. *Econometrica: Journal of the Econometric Society, 1,* 1–18.

Bicchieri, C. (2006). *The grammar of society: The nature and dynamics of social norms*. Cambridge University Press.

Bicchieri, C. (2017). *Norms in the wild: How to diagnose, measure, and change social norms*. Oxford University Press.

Brennan, G., Eriksson, L., Goodin, R. E., & Southwood, N. (2013). *Explaining norms*. Oxford University Press.

Centola, D., Willer, R., & Macy, M. (2005). The Emperor's dilemma: A computational model of self-enforcing norms. *American Journal of Sociology, 110*(4), 1009–1040.

Crawford, S., & Ostrom, E. (1995). A grammar of institutions. *American Political Science Review, 89*(3), 582–600.

Fisman, R., & Miguel, E. (2007). Corruption, norms, and legal enforcement: Evidence from diplomatic parking tickets. *Journal of Political Economy, 115*(6), 1020–1048.

Fudenberg, D., & Maskin, E. (1986). The Folk theorem in repeated games with discounting or with incomplete information. *Econometrica*, *54*(3), 533–554. https://doi.org/10.2307/1911307.

Gächter, S., & Schulz, J. F. (2016). Intrinsic honesty and the prevalence of rule violations across societies. *Nature, 531*(7595), 496–499.

Gelfand, M. J., Raver, J. L., Nishii, L., Leslie, L. M., Lun, J., Lim, B. C., Duan, L., Almaliach, A., Ang, S., Arnadottir, J., & Aycan, Z. (2011). Differences between tight and loose cultures: A 33-nation study. *Science, 332*(6033), 1100–1104.

Gintis, H. (2009). *The bounds of reason: Game Theory and the unification of the behavioral sciences*. Princeton University Press.

Gintis, H. (2010). Social norms as choreography. *Politics, Philosophy & Economics, 9*(3), 251–264.

Guala, F. (2013). The normativity of Lewis conventions. *Synthese, 190*(15), 3107–3122.

Guala, F. (2016). *Understanding institutions: The science and philosophy of living together*. Princeton University Press.

Guala, F., & Hindriks, F. (2014). A unified social ontology. *The Philosophical Quarterly, 65*(259), 177–201.

Guiso, L., Herrera, H., & Morelli, M. (2016). Cultural differences and institutional integration. *Journal of International Economics, 99,* S97–S113.

Hart, H. L. A. (1961). *The concept of law*. Clarendon Press.

Hindriks, F., & Guala, F. (2015). Institutions, rules, and equilibria: A unified theory. *Journal of Institutional Economics, 11*(3), 459–480.

Kaufmann, D., Kraay, A., & Mastruzzi, M. (2006). *Governance matters V: Aggregate and individual governance indicators for 1996–2005* (vol. 4012). World Bank.

Knack, S., & Keefer, P. (1997). Does social capital have an economic payoff? A cross-country investigation. *The Quarterly Journal of Economics, 112*(4), 1251–1288.

Kuran, T. (1995). *Private truths, public lies: the social consequences of preference falsification*. Harvard University Press.

Lewis, D. (1969). *Convention: A philosophical study*. Harvard University Press.

Mauro, P. (2004). The persistence of corruption and slow economic growth. *IMF Staff Papers, 51*(1), 1–18.

North, D. C. (1990). *Institutions, institutional change and economic performance*. Cambridge University Press.

North, D. C. (1991). Institutions. *Journal of Economic Perspectives, 5*(1), 97–112.

Paluck, E. L. (2009). Reducing intergroup prejudice and conflict using the media: A field experiment in Rwanda. *Journal of Personality and Social Psychology, 96*(3), 574–587.

Schotter, A. (1981). *The economic theory of social institutions*. Cambridge University Press.

Sillari, G. (2013). Rule-following as coordination: A game-theoretic approach. *Synthese, 190*(5), 871–890.

Vanderschraaf, P. (1995). Convention as correlated equilibrium. *Erkenntnis, 42*(1), 65–87.

Vanderschraaf, P. (1998). Knowledge, equilibrium and convention. *Erkenntnis, 49*(3), 337–369.
Vanderschraaf, P. (2001). *Learning and coordination: Inductive deliberation, equilibrium and convention*. Routledge.
Wittgenstein, L. (1953). *Philosophical investigations*. Blackwell.