University of Groningen

Norms that Make a Difference

Hindriks, Frank

Published in:
Analyse & Kritik

DOI:
[10.1515/auk-2019-410109](https://doi.org/10.1515/auk-2019-410109)

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2019

[Link to publication in University of Groningen/UMCG research database](#)

Citation for published version (APA):

Hindriks, F. (2019). Norms that Make a Difference: Social Practices and Institutions. *Analyse & Kritik*, 41(1), 125-145. <https://doi.org/10.1515/auk-2019-410109>

Copyright

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

Take-down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Frank Hindriks*

Norms that Make a Difference: Social Practices and Institutions

<https://doi.org/10.1515/auk-2019-0008>

Abstract: Institutions are norm-governed social practices, or so I propose. But what does it mean for a norm to govern a social practice? Theories that analyze institutions as equilibria equate norms with sanctions and model them as costs. The idea is that the sanctions change preferences and thereby behavior. This view fails to capture the fact that people are often motivated by social norms as such, when they regard them as legitimate. I argue that, in order for a social norm to be perceived as legitimate, agents have to acknowledge reasons for conforming to it other than the sanctions they might incur for violating it. In light of this, I defend a theory of institutions that does not only invoke equilibria, but also normative rules that are supported by normative expectations and, in some cases, normative beliefs.

Keywords: correlated equilibrium, institution, normative belief, normative expectation, rules-in-equilibrium, social norm, social practice

1 Introduction

It is sometimes said that there is honor among thieves. The intended meaning of this proverb is that even the unethical and disreputable adhere to common codes of conduct. If there really is honor among thieves, then the social practice of stealing is governed by norms (in addition to those involved in property rights). Thieves might, for instance, respect each other's territories. Consider police officers next. They have certain rights and obligations, such as to carry a gun and to fight crime. This reveals that also the practice of catching thieves is governed by norms, in this case formal rules. Thieves and police officers, as well as stealing and placing someone under arrest, are institutional phenomena. And they feature norms. But why is this? What roles do norms play in institutions?

According to equilibrium theories, institutions consists of strategies for interaction that are stable in that no one benefits from unilaterally changing her behavior. Think, for instance, of driving on the left-hand side of the road. When

*Corresponding author: Frank Hindriks, Dept. of Ethics, Social and Political Philosophy, Faculty of Philosophy, University of Groningen, e-mail: F.A.Hindriks@rug.nl

everybody else does it, doing so is the best course of action for you as well—in that it maximizes your expected utility. As this also holds for the others, the practice is mutually beneficial. To be sure, institutions can be harmful, as when cooperation breaks down and gives rise to a vendetta. However, even in such cases an individual still does not stand to benefit from unilaterally switching to a different course of action. Thus, institutions generate collective benefits. Furthermore, those collective benefits explain why institutions exist and persist (Hindriks/Guala, forthcoming). It follows that generating such benefits is the function of institutions (Schotter 1981; Tuomela 2013; Hindriks/Guala, forthcoming). In this paper, I argue that institutions feature social norms that contribute to the performance of this function.

Against this background, I propose that institutions are norm-governed social practices. Social norms can affect behaviors because of sanctions people might incur for violating them. According to equilibrium theories, sanctions change people's preferences in cooperation games such that they cooperate (Ullmann-Margalit 1977; Schotter 1981). Consider two farmers who agree to help each other harvest their crops. Once the other has helped the one, she might wonder why she would still help the other. She might do so because the other will disapprove of her if she does not return the favor. However, there is a second way in which norms can affect behavior. People might regard a norm as legitimate and take themselves to have good reasons for conforming to it (Bicchieri 2006; Brennan et al. 2013). Presumably, this explains why those who violate norms typically struggle with their decision to do so (Bandura 2016). The point to appreciate is that there is often more to the decision to conform than a cost-benefit calculation.

The main challenge I address in this paper is how to incorporate the notion of (perceived) legitimacy in an equilibrium theory of institutions. To this end, I propose to enrich equilibrium theories with the notion of a normative rule, which is a rule that specifies what people are supposed to do. In order for it to be in force, people must expect that others believe the rule applies to them. When someone regards a social norm as legitimate, she believes this herself. Furthermore, this belief is in part supported by her (justified) expectation about what others believe. Finally, I argue that the perceived legitimacy of a social norm can make a difference to how people interact. In doing so, I defend and develop the Rules-in-Equilibrium theory (*RiE*) that Francesco Guala and I introduced elsewhere (Hindriks/Guala 2015; Guala/Hindriks 2015). The core idea of that theory is that the notion of an equilibrium does not suffice for adequately analyzing institutions, but needs to be complemented with that of a rule. Guala and I have thus far focused on the signaling or coordinating role that rules can play. Here I attend to how rules can shed on the normative dimension of institutions, which serves to promote cooperation.

In *section 2*, I introduce the equilibrium approach to social practices, norms and institutions. In *section 3*, I argue that there can be social practices that do not involve social norms. And in *section 4*, I explain what it means for a social norm to govern a social practice. I build up to this by first explicating what it means for a social norm to exist, to regard it as legitimate and to follow.¹

2 Institutions and Social Norms

Institutions involve patterns of behavior that give rise to mutually beneficial outcomes. Such patterns of behavior are social practices. Paradigmatic examples include marriage, money and property. Strikingly, each of these involves norms. For instance, depending on context, marriage is supposed to be monogamous. Money and property are closely related institutions. To own something is in part to have the right to its use. Money serves to acquire property rights over what is bought. Thus, it involves a second-order norm or right to change first-order property rights. These examples suggest that, in addition to social practices, social norms form an important component of institutions.

The norms of institutions either coordinate people's behaviors or enable them to cooperate. Many coordination norms are of the form: first-come-first serve. Examples include norms that govern standing in line and first-occupation rules of property. In contrast, a number of cooperation norms are of the form: I-will-scratch-your-back-if-you-scratch-mine. Think, for instance, of farmers who help each other to harvest their crops. As another example, consider crew members who struck deals between themselves that they would deliver only light lashes with a whip—to merely 'scratch' the offender's back—when they were ordered to punish someone. They did so to ensure they were treated the same should they ever found themselves on the receiving end at some time in the future. I argue that coordination norms stabilize coordination and cooperation norms enable cooperation. They thereby secure or increase the collective benefits that a social practice generates.

In this section, I discuss what role social norms play in the equilibrium approach to institutions. As will become apparent, initially they played no role in

¹ This is the first of three papers about institutions and norms. In the second part of this trilogy, I explain what it means for an institution to be stronger than another one, building on the account of motivation for conformity that I provide here. In the third part, I defend the idea that a theory of institutions can and should integrate the behavioral, normative and symbolic dimensions of institutions, all of which feature in this paper as well.

equilibrium theories. However, later they were explicated in terms of sanctions and thereby modeled as costs. According to the equilibrium approach, institutions can be modeled as equilibria in repeated games. An equilibrium is a combination of strategies or course of action that is stable in that nobody has an incentive to unilaterally deviate from it. Suppose that you and I are in Berlin. I want to meet you, but I do not care about where we meet. In contrast, you do not particularly care about meeting me, but you want to see the Brandenburger Tor, come what may. Then each of us has an overriding incentive to go there. This combination of strategies forms an equilibrium. Although this is a rather specific case, many situations reoccur over time. People will then rely on the same strategies, which means that a behavioral regularity ensues. The equilibrium approach equates institutions with such behavioral regularities.

The players or participants in a game have interests that either align or conflict: they align in coordination games (Lewis 1969; Schotter 1981); they conflict in cooperation games (Ullmann-Margalit 1977; Schotter 1981). Institutions can be regarded as solutions to coordination or cooperation problems. Coordination games have two equally good solutions that serve the interests of all players equally well, or at least more or less so. The paradigm example is driving on a particular side of the road, for instance left in the UK and right in continental Europe. In a repeated coordination game, players encounter each other in a recurrent situation. They have to decide what to do or select a strategy. The behavior that people will in fact display is a convention. The problem is, however, that there is as such no basis for preferring one option to the other.

David Lewis (1969) has argued that people choose on the basis of precedence. If they observe a particular pattern of behavior—say driving on the left—then this is what they expect others to do. And giving this expectation, they prefer to do so as well. More precisely, they have conditional preferences: they prefer a particular behavior conditional on what they expect others to do. Given such conditional preferences, the members of the relevant population succeed in coordinating their behaviors. Lewis argued that this will work only if all of this is out in the open (or more technically common knowledge). The resulting strategy profile is stable in that no one has an incentive to deviate unilaterally. For instance, once people have converged on a particular side of the road, nobody has an incentive to drive on the other side. As a consequence, everybody benefits.

Lewis' account of conventions consists of four conditions. Consider a situation *S* that can be represented as a coordination game, which means that people's payoffs converge. A behavioral regularity is a convention among a population *P* in *S* exactly if the following conditions are met: (1) The members of *P* conform to the regularity; (2) They expect others to conform; (3) They prefer to conform con-

ditional on others doing so; (4) All of this is common knowledge.² In this way, the equilibrium approach captures the behavioral dimension of institutions. As I discuss in section 2.1, Lewis regards conventions as norms. Strikingly, however, he does not incorporate norms in his model. This raises the question whether and how the equilibrium approach can do justice to the normative dimension of institutions. Part of the answer is provided by proponents of the equilibrium approach who have modeled institutions that solve cooperation games.

In cooperation games, the interests of the participants conflict. They face a choice between cooperating and not cooperating or defecting. Everybody would benefit from uniform cooperation. But each of them stands to gain from unilateral defection. This means that it is (individually) rational to defect. Because of this, none of the participants cooperates and the outcome that materializes is suboptimal. I might consider to help you move house on the condition that you will help me at some later point in time. But as you have no incentive to help me then, I do not help you now. This in spite of the fact that we both realize that everybody would benefit from cooperating.

This problem can be solved by means of sanctions that align the interests of the participants. As mentioned, sanctions can be formal or informal. Formal sanctions include fines and imprisonment; informal sanctions encompass various forms of disapproval (Ullmann-Margalit 1977; Schotter 1981). Sanctions can be seen as costs that affect the preferences agents have. Participants can incur these costs when they do not conform to an institution. Sanctions affect behavior because the prospect of these costs can change the preferences of the participants. Whether they do so, depends on how high these costs are and how probable it is that non-conforming behaviors are detected. When the expected costs are high enough, the game changes such that participants cooperate (Crawford 1995).³

Thus, the norm of an institution can be a coordination norm or a cooperation norm that is either formal or informal. Focusing on formal norms of cooperation, Andrew Schotter observes that ‘the famous prisoners’ dilemma game can be solved through the use of a binding contract that is enforceable by an external authority’ (Schotter 1981, 11). The prospect of a fine or of imprisonment removes the incentive to break formal rules. Edna Ullmann-Margalit (1977) develops basically the same argument for informal norms of cooperation. She argues that social pressure, exerted by means of some form of disapproval, can be seen as a sanction

² Bicchieri 2006, 31–42, defends a similar analysis of conventions that does not require common knowledge.

³ Crawford/Ostrom 1995 model sanctions as delta parameters that change the payoffs of the relevant sanctions. They can increase the costs of not cooperating such that agents prefer to cooperate.

or cost that participants can incur. The upshot is that both Schotter and Ullmann-Margalit effectively model norms as costs. The key claim, then, is that institutions are maintained in part because norms come with such costs.⁴

Andrew Schotter (1981) and Raimo Tuomela (2013) have proposed that the function of social practices and institutions is to generate collective benefits. Recently, Francesco Guala and I have defended this idea arguing that those collective benefits explain why institutions exist and persist (Hindriks/Guala, forthcoming). Cooperation norms do so because they enable cooperation. Although it is in principle possible for cooperation to arise in the absence of such norms, the incentives that sanctions provide can be essential to the existence of an institution. When they are, they enable the institution to perform its function.

In contrast, conventions are self-reinforcing. Because of this, sanctions are not needed for creating and maintaining a mutually beneficial behavioral regularity. It follows that conventions need not involve norms (Bicchieri 2006). However, conventions are rarely if ever exceptionless—both Lewis (1969, 76–80) and Bicchieri (2006, 38) allow for conventions to which not all members of the population conform. And some are more fragile than others. Because of this, norms do have a role to play with respect to conventions as well. As norms make conforming more attractive, they serve to stabilize conventional patterns of behavior. They reinforce existing ways of coordinating and thereby secure the benefits that conventions generate. Given that the function of conventions is also to generate collective benefits, coordination norms serve to prevent that conventions malfunction. In these ways, social norms make a difference to how people interact.

Thus, the function of institutions is to generate collective benefits (Hindriks/Guala forthcoming). They perform this function in part because of the social norms they feature. Such norms, in turn, serve to enable cooperation or stabilize coordination.

3 Social Practices, Normative Rules and Signaling Rules

Institutions are norm-governed social practices. In order to defend this claim, I need to explain what social practices are. Social practices involve interdependent

⁴ Even if they are collectively beneficial in the sense specified, institutions can have downsides (Knight 1992). For instance, Ullmann-Margalit 1977 and Schotter 1981 discuss how institutions such as property rights and inheritance laws preserve inequalities.

behavioral regularities. Those regularities are interdependent in that what one participant prefers to do depends on what others do or are expected to do.⁵ The simplest conception of social practices is that they are nothing but interdependent behavioral regularities. This can suitably be captured by the equilibrium approach. They would be equilibria in repeated games. This proposal faces two challenges. The first is that social practices have a symbolic dimension. The claim is that all social practices exist due to symbols that facilitate coordination or enable cooperation. Think, for instance, of traffic signs and traffic lights that help people to travel efficiently and safely. Such symbols are signaling devices. As I discuss below, they feature in signaling rules. The second challenge is that social practices have a normative dimension. The claim is that all social practices involve norms or normative rules, just as in the cooperative farmers example mentioned above.

Social practices typically involve rules, or so I propose. This is important because it means that a general theory of social practices must be hybrid and feature both equilibria and rules. Elsewhere, Guala and I introduced the Rules-in-Equilibrium theory (*RiE*), which invokes signaling rules (Hindriks/Guala 2015; Guala/Hindriks 2015). In *section 3.1*, I argue that it applies to conventions (as well as to coordination games with asymmetric payoffs). As signaling devices serve to coordinate behavior, the theory does not extend to cooperation games. In *section 3.2*, I argue that conventions need not involve normative rules, because they are self-reinforcing. In contrast, cooperative practices tend to be unstable if not supported by normative rules. Because of this, I propose in *section 3.3* that such practices usually involve social norms.

The Rules-in-Equilibrium theory (*RiE*) allows for normative rules. However, it models them as costs. This is adequate insofar as conformity to norms is due to sanctions. However, it is unsatisfactory to the extent that people do so because they perceive norms as legitimate. It is this idea that I set out to capture in *section 4*. The main point here is that conventional practices typically involve signaling rules and cooperative practices normative rules, while some involve both.⁶

5 In a similar vein, Brennan et al. 2013, 16, define social practices as behavioral regularities that are explained by desires about the relevant behavior (or beliefs about such desires) as well as by the fact that those desires are known. The only substantial difference between our views is that I require belief rather than knowledge. People could, for instance, sustain a practice based on outdated or otherwise mistaken beliefs about each other's preferences.

6 Searle 1995; 2010 proposes that institutions are systems of constitutive rules. Hindriks/Guala 2015 and Guala/Hindriks 2015 argue that this is consistent with the Rules-in-Equilibrium theory. See Hindriks 2009; 2013a, for more on constitutive rules.

3.1 Conventions and Signaling Rules

According to the Rules-in-Equilibrium theory (*RiE*), social practices coordinate behaviors by means of signaling devices (Hindriks/Guala 2015; Guala/Hindriks 2015). Devices that fulfill this signaling role encompass entities as diverse as wedding rings, military uniforms and traffic lights. They introduce new strategies for action that are mutually beneficial (Hindriks 2015, 465). People might, for instance, wear a wedding ring in order to signal that they are married and that making advances is not appreciated. Although soldiers wear uniforms as camouflage during combat, they also serve to make their role salient in other contexts. Signaling devices introduce new strategies that are conditional on their presence. They expand the set of available strategies by facilitating people to act differently when they are there. This in turn enables them to align or coordinate their behaviors in ways that are beneficial to all. In this way, they give rise to conventions.

Let me use the example of traffic lights to make this precise. Consider an intersection with cars converging from several sides. Ordinarily, drivers can choose between two strategies: ‘stop’, and ‘go’. Their structure is simply ‘do *A*’. Traffic lights facilitate strategies that have a more complex structure, such as: ‘If red, stop.’ Because of the signal that features in it, this is a signaling rule. Using ‘*D*’ for ‘signaling device’, the structure of such a rule is: ‘If *D*, do *A*’. Because of their conditional structure, signaling rules serve to make traffic safer and more efficient. In this way, signaling or correlating devices correlate or coordinate behaviors so as to give rise to so-called correlated equilibria (Aumann 1974; Vanderschraaf 1995; Gintis 2007).

Thus, *RiE* combines the notion of an equilibrium with that of a signaling rule in order to explain coordination (Hindriks/Guala 2015; Guala/Hindriks 2015). The core claim is that interdependent behavioral regularities in coordination games arise in response to signaling devices. And such devices feature in signaling rules that present as the thing to do for people to conditionalize their behavior on some signal. That signal serves to coordinate their behavior such that they engage in mutually beneficial interaction and secure collective benefits. The upshot is that not all social practices can be explained in terms of equilibria only, and that a hybrid theory that also invokes rules is needed to account for all of them.

3.2 Conventions without Normative Rules

Conventions need not involve normative rules. The reason for this is that they are self-reinforcing. Lewis’ (1969) model of conventions, which I discussed in *section*

2, does not invoke normative rules. Furthermore, he claims that ‘convention’ is not a normative term (1969, 97). By this he means that conventions are not intrinsically normative phenomena. Christina Bicchieri (2006) concludes from this that there can be conventions without norms (see also Brennan et al. 2013 and Southwood 2018). Strikingly, however, Lewis also claims that conventions are “a species of norms” (1969, 98). He observes that conforming to a convention is in the interests of all participants. As this is common knowledge, each agent knows that conforming is in their own interest as well as those of others. Lewis argues that, because of this, “there is some presumption that one ought to conform” (1969, 99). And he concludes that, necessarily, when some behavior is conventional, it will also be obligatory.⁷

However, the fact that conventional behavior is in everybody’s interests does not as such establish that there is an obligation to conform. That it is in *my* interests merely means that it is instrumentally rational for me to conform. And it cannot be taken for granted that the fact that my conforming is of interest to *others* obligates me in any way. Brennan et al. (2013) argue that conventions do not necessarily obligate. They imagine a group of friends who loathe rules but support a convention all the same. They develop a regular way of meeting that “has to be wholly optional on all sides” (Brennan 2013, 18). As they dislike rules, the friends do not accept any norms. Hence, their regular meeting is a convention without a norm.

At the same time, it appears that many conventions are in fact governed by norms. Brennan et al. (2013, 19) argue, for instance, that conventions can become normative because they protect or promote some value. This even holds for a simple convention such as driving on the left-hand side of the road. It is a norm, and even a law, because of safety concerns. Robert Sugden argues that conventions create expectations of conformity that are legitimate and, as such, ought to be honored: “We feel entitled to expect others to follow conventions when they deal with us, and we recognize that they are entitled to expect the same of us.” (1986, 154)

When interpreted as a tendency claim, this seems right. Francesco Guala and Luigi Mittone (2010) have found some initial empirical support for this claim. They conclude that “mere repetition of a collective task does enhance the conformity to a convention” and that “conventions have a tendency to become social norms, acquiring an intrinsic normative power” (2010, 754–755). The point can be made

⁷ Gilbert criticizes Lewis for not treating conventions as intrinsically normative. She argues that conventions are quasi-agreements (Gilbert 1989, 369 and 350). As such, they are conceptually linked to joint intentions, which she takes to generate reasons for conformity (see Hindriks 2013b for a discussion).

in terms of a distinction between descriptive and normative conventions. The former are mere social practices, the latter involve (coordination) norms and are as such institutions. In these terms, the claim is that, even though descriptive conventions tend to develop into normative conventions, it is possible for there to be conventional practices without norms.

The idea that there could be social practices without norms is sometimes met with incredulity. The underlying concern appears to be that normativity permeates social interaction. At this point, it is important to distinguish between different kinds of normative standards: deontic and evaluative. Social norms are deontic standards in that they feature obligations. They specify what is right or wrong. However, social practices can also involve internal standards as to what a good way of performing an activity is. Just as there are better and worse ways of playing backgammon, there are, for instance, better or worse ways of being a host. Due to such standards, certain moves are more appropriate than others. An experienced host will have a better sense of this than an inexperienced host. None of this means, however, that particular moves are mandated. Consider hosting a party. Some people like doing it and are good at it. But in many contexts, it is acceptable not to throw a party. Even so, there are standards concerning how to organize one. This reveals that *there can be better or worse without right or wrong*.

Thus, the view that normativity permeates social interaction is consistent with the claim that there can be social practices without social norms. The idea is that, even though many social practices feature both deontic and evaluative standards, there can be social practices that involve evaluative standards only. Even so, there might be social practices that feature neither deontic nor evaluative standards. Recall the group that loathes rules. It might also dislike evaluative standards. The upshot is that conventions involve signaling rules, but need not feature normative rules. This means that there can be social practices that are not institutions. In *section 4*, I give further content to the claim that institutions are norm-governed social practices. For this purpose, it is useful to know that social practices and social norms are distinct phenomena.

3.3 Cooperative Practices with Normative Rules

Cooperative practices almost always involve social norms. The reason for this is that such practices tend to be unstable if not supported by normative rules. Peyton Young defines norms as equilibria: a norm is “an equilibrium behavior in a game played repeatedly by many different individuals in a society where the behavior is known to be customary” (Young 1998, 16). However, social norms differ from customs. They obligate certain actions. More specifically, social norms in-

volve obligations that feature in normative rules. It follows that a theory of social norms should invoke normative rules in addition to equilibria.

As a first approximation, I will take social norms to be generally accepted normative rules, or more precisely normative rules of which it is generally known that they are accepted (Brennan 2013, 2–4 and 172). Such rules feature prescriptions, prohibitions or permissions (or second-order rights or obligations). For convenience, I focus on prescriptive norms the rules of which present an action as obligatory. Against this background, the question becomes whether there can be interdependent behavioral regularities in the absence of generally accepted normative rules.

In cooperation games, the interests of the participants conflict. This means that in one-shot versions of these games the best option players have is to defect. However, in indefinitely repeated games it can be instrumentally rational to cooperate. This means that cooperative behavior can in principle arise spontaneously and independently of any sanctions. It follows that cooperative practices do not have to involve cooperation norms. However, they often do so in practice. Cooperation norms make cooperation attractive in situations in which it was not before. As discussed in *section 2*, they enable and sustain cooperation.

The players in cooperation games do not always end up cooperating. Strikingly, however, even uncooperative equilibria are often governed by norms. The thing to realize is that uncooperative practices can originate from cooperative ones. Consider vendettas. They typically arise between groups who cooperated for a prolonged period of time. However, this pattern of cooperation broke down for some reason resulting in multiple rounds of retaliation. In such cases, cooperation between groups was governed by a social norm that proscribed sanctions for violations. Since cooperation has broken down, participants now incur these sanctions. People retaliate in response, thus setting in motion a harmful pattern of behavior. It might be that this process is strengthened, for instance by considerations of honor (Brennan et al. 2013).

Finally, there can be cooperation games that never give rise to cooperation. This can be illustrated in terms of the prisoner's dilemma. The district attorney separates the two prisoners hoping that she can get them to inform on each other. She might do so repeatedly with a substantial number of suspected felons. If so, the activity of interrogating them separately turns into a social practice, this in spite of the fact that the people interrogated end up defecting time and again. The upshot is that, in contrast to uncooperative practices, cooperative practices are almost always governed by social norms.

Thus, social practices are interdependent behavioral regularities that are almost always governed by rules, be it signaling rules, normative rules or both. The Rules-in-Equilibrium theory (*RiE*) features signaling rules. Furthermore, it allows

for normative rules. However, it models them as costs. This is adequate insofar as conformity to norms is due to sanctions. However, it is unsatisfactory to the extent that people do so because they perceive norms as legitimate. It is this idea that I set out to capture in *section 4*. This also serves to extend the theory to cooperation games. And it is a crucial step in explicating the claim that institutions are norm-governed social practices.

4 Institutions as Norm-Governed Social Practices

Social practices and social norms are the two faces of institutions. Having argued that there can be social practices without social norms, I now turn to the question of how exactly they relate to one another when they constitute an institution. To this end, I discuss the conditions under which a social norm exists (*section 4.1*). Subsequently, I explicate what it means for a social norm to be regarded as legitimate (*section 4.2*), for it to be legitimate and for an agent to follow it (*section 4.3*). This serves to develop the Rules-in-Equilibrium theory (*RiE*) and to complete my defense of the thesis that institutions are norm-governed social practices (*section 4.4*).

4.1 What Is a Social Norm?

In order for a social norm to govern a social practice, the norm has to exist. This will only be the case if its normative rule features in people's attitudes. A natural suggestion is that a social norm exists exactly if it is generally known that people believe its normative rule applies. By this, I mean to say that they believe that the relevant activity is obligatory for those who participate in the social practice. In other words, the claim is that those who encounter a situation that constitutes a coordination game or a cooperation games are believed to be obligated to act in a particular manner and that this is generally known. This 'normative-beliefs view', as I call it, is a version of the theory that Brennan et al. (2013, 2–4 and 172) defend when they characterize social norms as normative principles or rules of which it is generally known that they are generally accepted. In terms of beliefs, the claim is simply that social norms are generally known normative beliefs.⁸

⁸ Brennan et al. 2013 allow for explications of acceptance in terms of attitudes other than beliefs, such as desires.

A phenomenon known as pluralistic ignorance reveals that this is too strong (Prentice/Miller 1993). The famous example concerns students at some college who engage in heavy drinking in the weekend. Each does so because she believes she is supposed to. By this I mean that they believe that the others believe they ought to. Furthermore, this matches how others behave. Many students drink heavily on weekends. And they criticize those who do not. At the same time, however, each of them dislikes drinking heavily. They are discouraged from sharing their dislikes, because they believe doing so would be disapproved of. After all, drinking is the norm. In such cases of pluralistic ignorance, participants conform only because of reasons that are external to the norm, the disapproval they would incur for not doing so or the approval they meet when they do. The upshot is that a social norm can exist in the absence of normative beliefs.

The possibility of pluralistic ignorance reveals that all that is required for a social norm to exist is that people believe that they are supposed to behave in a certain way. They do not need to believe that they ought to do so. This can be explicated in terms of normative expectations. As I analyze the notion, a normative expectation is an expectation that others believe that the normative rule applies. Someone who possesses a normative expectation expects others to believe that everybody ought to behave in the manner specified by that rule. In light of this, I propose the following analysis of the existence of a social norm [E]:

[E] A social norm *exists* in a population exactly if its normative rule features in normative expectations of a substantial number of its members.

I refer to it as ‘the normative-expectations view’. It differs from the normative-beliefs view in two respects. First, the normative-expectations view does not require that the relevant individuals accept the norm (in the sense that they believe that it applies). All that is needed is that they acknowledge it (by which I mean that they believe they are supposed to conform to it). Second, in contrast to the normative-beliefs view, it does not require that people know about the attitudes that the others have. Expecting that they have a belief suffices. The normative-expectations view allows for normative rules to feature in first-order beliefs. What is distinctive of it is that it does not require this.

Bicchieri defends a theory that is similar to the normative-expectations view in many respects. In spite of this, there are three important differences between the two. First, Bicchieri restricts the notion of a social norm to normative rules that solve cooperation problems, whereas I extend it to coordination problems. Second, she relies on a conception of normative expectations on which they are, strictly speaking, not normative. As she defines the notion, normative expectations are higher-order empirical expectations (Bicchieri 2006, 11). More precisely,

someone possesses a normative expectation exactly if she expects others to expect her to conform to a descriptive rule of the form ‘Everybody do A’. To be sure, Bicchieri suggests in her informal discussion that they do involve obligations. Even so, the expectation that someone conform to a descriptive rule differs from the belief that a normative rule applies. The former is an expectation about someone’s behavior, the latter a belief about what someone ought to do.

Third, Bicchieri explicates the existence of a social norm in terms of preferences. A norm of cooperation exists exactly if a sufficient number of participants prefer to conform to it conditional on empirical and normative expectations (Bicchieri 2006, 11).⁹ Because of this, I refer to her view as ‘the conditional-preferences view’. Nick Southwood (2018) criticizes this view by pointing out that people might secretly prefer to violate the norms that are in force conditional on others acting in this way and expecting each other to do so. As an example, he mentions considers a highly repressive society in which people have a conditional preference for acting according to libertine principles. In this society, people do not conform to those principles and do not expect others to do so. Hence, the libertine principles are not norms in the society at issue. However, as people have conditional preferences, Bicchieri’s analysis implies that they are. This reveals that conditional preferences do not suffice for there to be social norms. Normative expectations are required as well.

The normative-expectations view also requires conditional preferences. Normative expectations feature normative rules. Such rules concern situations that constitute coordination problems or cooperation problems. In such situations, people prefer to coordinate or cooperate on the condition that others do so too and are supposed to do so—that is, in effect, conditional on empirical and normative expectations. The difference is that the normative-expectations view does not regard conditional preferences as sufficient, but requires normative expectations as well.

4.2 The Perceived Legitimacy of Social Norms

Albert Bandura (2016) observes that those who violate a norm often struggle with doing so. This makes little sense if conformity is motivated merely by sanctions. These simply add to the costs or benefits of the available options. They give rise

⁹ In order for a social norm to give rise to a self-sustaining regularity, there have to be sufficiently many followers or agents who have unconditional preferences of conformity: “A social norm is followed if the set of its followers is sufficiently large.” (Bicchieri 2016, 12) An agent unconditionally prefers to conform only if he believes enough others in fact conform.

to a cost-benefit calculation rather than a struggle. However, people often regard social norms, along with the associated sanctions, as legitimate (Bicchieri 2006). This explains the struggle, because it puts pressure on people to justify to themselves when they are tempted to violate a norm. In this way, it makes sense of the fact that norm violators are often conflicted.

As discussed in *section 2*, equilibrium theories of social norms reduce them to sanctions. They influence behavior if and because they increase the expected costs of violating those norms. Now, suppose that a social norm is in equilibrium. Then everybody acts according to that norm. If there are exceptions, this must be because the agents who violate the norm fail to meet one of the conditions that constitute the equilibrium. In particular, they might exhibit atypical preferences. A rich person, for instance, might not regard a fine as costly; and an insensitive person might care little about a disapproving frown. Such people make a cost-benefit calculation that differs from that of the majority. The notion that social norms can be perceived as legitimate plays no role in equilibrium theories.

Bicchieri (2006) argues that social norms are often perceived as legitimate, and, what is more, that this also motivates people to conform. When people perceive a norm as legitimate, they believe that there are good reasons for the norm to be in place. More specifically, they actually believe that the norm captures reasons that are strong enough to obligate them. Furthermore, the belief that there are such reasons can motivate them. They might have a preference for acting on good reasons or, more precisely, on the basis of justified beliefs about what they ought to do.

But how can this notion of perceived legitimacy be incorporated in a theory of institutions? Consider what the normative-beliefs view and the conditional-preferences view say about this. Inspired by H. L. A. Hart, Brennan et al. (2013) take the fact that people have generally known normative beliefs to imply that they regard the relevant normative rule as legitimate. They claim that social norms “bootstrap their own authority” (2013, 6, see also 42). This suggests that normative beliefs are self-justifying, or at least that people regard them as such. Such beliefs derive support from the very existence of a behavioral regularity that corresponds to the relevant normative rule (2013, 69). Crucially, the fact that an agent perceives a norm as legitimate can motivate him to conform to it. He will take himself to have reasons for following it. Those reasons are internal or intrinsic to the norm and as such independent of any sanctions. They are what I call ‘norm-internal reasons’.

According to the conditional-preferences view, perceiving a social norm as legitimate is a matter of regarding the relevant normative expectations as well-founded (Bicchieri 2006, 15 and 23–24). Because normative expectations can influence people’s preferences, the fact that someone perceives a norm as legitimate can provide him with motivate to act accordingly (Bicchieri 2006, 11; 2016,

35). Bicchieri distinguishes two ways in which normative expectations can influence someone's motivation. The first concerns sanctions. An agent might want to conform to a norm at least in part because others expect her to do so. Someone might seek approval for meeting other people's expectations or avoid disapproval for frustrating them. Alternatively, it could be because she perceives the normative rule as legitimate. The idea is that her normative expectation influences her preferences because she accepts it as well-founded or justified.

Thus, the normative-beliefs view and the conditional-preferences view have different conceptions of what is required for people to perceive a normative rule as legitimate. According to the former, all it takes is for them to possess normative beliefs that feature the rule: normative beliefs are regarded as self-justifying. In contrast, the latter explicates it in terms of them normative expectations that are regarded as justified.

Neither of these two views is fully adequate. The normative-beliefs view, according to which social norms are self-justifying, entails that people tend to regard any social norm as legitimate. However, the phenomenon of pluralistic ignorance reveals that there is a whole range of social norms for which this does not hold.¹⁰ According to the normative-expectations view, in contrast, it is a contingent matter whether a social norm is perceived as legitimate. However, it raises the question why perceived legitimacy would, in the first instance, be a matter of justified normative expectations, rather than justified normative beliefs. Someone who regards a normative expectation as justified will be inclined to form the corresponding normative belief. It seems, however, that this belief lies at the core of what it means for a social norm to be perceived as legitimate. Arguably, such a belief will be suitably justified only if the corresponding empirical and normative expectations are justified. The conditional-preferences view fails to capture these intricacies of perceived legitimacy. In contrast, the normative-expectations view explicates it in terms of normative beliefs that an agent happens to regard as justified for the reasons just mentioned. The upshot is that an agent can perceive a social norm as legitimate and that this can motivate her to conform. In this way, the social norms that feature in institutions can as such make a difference to behavior.

¹⁰ This means that people need not regard a social norm as legitimate in order to sanction violations, and that they may sanction others even if they themselves are not sufficiently motivated to conform.

4.3 Following and Violating a Norm

Having discussed what it takes for someone to *perceive* a social norm as legitimate, I continue to ask what it means for a social norm to *be* legitimate or, as I will also say, to have authority. According to the normative-expectations view, a social norm has authority exactly if the normative beliefs people have, are in fact suitably justified. This requires that they are supported by empirical and normative expectations that are themselves not only justified but also true, which means that there is in fact an (approximately) corresponding regularity and others do in fact believe that the normative rule applies. An appreciation of the authority of a social norm is vital for understanding what it means to follow such a norm, or so I argue in this section.

Brennan et al. maintain that following norms is a matter of “doing what the norms require because of the norms” (2013, 10). This means that someone who follows a norm is motivated by the norm as such. I will refer to the relevant reasons, which were discussed in *section 4.2*, as ‘norm-internal reasons’. Thus, an agent follows a social norm exactly if: he conforms to it because of norm-internal reasons. If he is also motivated by sanctions, he would still conform if there were no prospect of sanctions. This means that he conforms to it because of the authority it has. Note that violating a norm does not require an agent to recognize norm-internal reasons. An agent violates a social norm exactly if she does not conform to it even though she correctly believes it exists.

On the face of it, the notion of following a norm is inconsistent with equilibrium theories of institutions. Such theories are concerned with the incentives people have and therefore with their motives, not with norm-internal reasons. I will argue, however, that this notion can be incorporated in the Rules-in-Equilibrium theory (*RiE*). In *section 4.1*, I suggested that people can prefer to act on good reasons or, more precisely, to act on the basis of justified beliefs about what they ought to do. When they do, those beliefs increase the agent’s payoffs for conforming to the norm. And they can do so to such an extent that they motivate the agent conform. Given the proposed explication of the authority of social norms, this means that an agent follows such a norm exactly if: he prefers to conform to it because of a normative belief that is justified by empirical and normative expectations that are both justified and true. This incentive-based explication matches the reason-based explication just presented. And it reveals that the notion of a norm-internal reason can be brought to bear on the explanation of behavior in a systematic manner by employing the explanatory framework that game theory provides.

4.4 Being Governed by a Norm

To complete my account of institutions, I now ask how a social norm can govern a social practice. A natural thought is that this can be explicated in terms of the account of following a norm just presented. The idea would be that for a social norm to govern a social practice is for its participants to follow that norm. There will indeed be an institution when this condition is satisfied. However, it will also be perceived as legitimate and it will in fact have authority. Even though it is important to understand what it means for an institution to be authoritative, the main question I am after is what an institution is. As there can be institutions that are not perceived as legitimate, the explication in terms of norm-following is too demanding. It would entail, for instance, that there is no use for sanctions, as norm-followers conform to a norm for norm-internal reasons. In practice, however, sanctions play important roles in formal as well as informal norms. Even if people are primarily motivated by norms as such, it can be that the prospect of sanctions tips the balance in favor of conformity. Consider someone who is tempted to cheat on someone even though she recognizes the significance of exclusivity for her relationship. Now, suppose that her commitment is not strong enough, but that she ultimately decides not to cheat because of the prospect of disapproval if others found out. It seems that the norm of exclusivity governs relationship practices even if there are many people who are just like her. The norm constrains an activity that is perceived as tempting. Hence, there is little reason to believe that all of those who participate in the institution of monogamy are motivated primarily by norm-internal reasons or by its intrinsic value.

An alternative proposal is that a norm governs a practice exactly if its participants, or a substantial number thereof, conform to it. This proposal is both too weak and too strong. It is too strong because an institution can exist even in the face of widespread violations. The institution of monogamy does not cease to exist, for instance, if many people cheat. It is too weak in that people might conform to a normative rule without being motivated by it. Think of the convention of driving on the right-hand side of the road. As discussed above, it is self-reinforcing. Now, suppose there is a social norm that mandates this behavior. It exists because people have the requisite normative expectations. However, it does not motivate anybody. In such a situation, the norm does not govern the practice. Instead, it is epiphenomenal. Hence, the convention does not constitute an institution. This reveals that a norm governs a practice only if it has an effect on the motivation of a substantial number of participants. It follows that norm-conformity is not sufficient for norm-governance. The upshot is that neither norm-following nor norm-conformity is necessary for norm-governance and norm-conformity is not sufficient.

The alternative that I propose is that a social norm has to exert a pull on people in order for it to govern a social practice. Consider the speed limit and suppose that everybody exceeds it. The speed limit applies to each and every one of those who speed. Furthermore, they can be fined for speeding. This reveals that, even though everybody violates it, there is some practice that is governed by the norm. But which practice? The norm specifies a limit or constraint on how fast someone is to drive on a public road. This provides for a description of the activity that leaves open how it is in fact performed. More generally, the social practice that is governed by a social norm is to be specified in terms that are neutral with respect to whether or how it is performed. Against this background, the notion of norm-governance can be explicated as follows [G]:

[G] A social norm *governs* a social practice exactly if a significant number of its participants are motivated to conform to its normative rule to a nontrivial degree.

In more colloquial terms, the proposal is that a norm governs a practice when people regard conforming to it as at least somewhat important. When this is the case, the social practice is an institution.

But how if at all is this consistent with the Rules-in-Equilibrium theory (*RiE*) of institutions? As I have developed it here, *RiE* takes social practices to be descriptive rules in equilibrium. The obvious way to extend this is to say that institutions are normative rules in equilibrium. However, if a rule is often violated, it is not in equilibrium. Consider the speeding example once again. It might be that the speed limit is set at X and that the equilibrium is $X + 10$. In such a situation, there is a mismatch between the norm and the equilibrium. The thing to realize is that, in the absence of the norm, the equilibrium might be $X + 20$. If this is the case, the norm affects their motivation. And, because speed is a matter of degree, it also influences their behavior, even if it falls short of conformity. This example suggests that institutions are normative rules that are at least approximately in equilibrium. However, things are different when violating a rule is not a matter of degree. Suppose that bikers ignore traffic lights on a massive scale. Then there is a discrepancy between the social practice and the social norm. And the relevant normative rule is not even approximately in equilibrium. To allow for this, I propose to retain *RiE* as a theory of social practices, but to refer to the account of institutions as ‘the *Rules-AND-Equilibrium* theory’ (*RaE*). The rule and the equilibrium are connected because, even though they diverge, the rule governs the practice.¹¹

¹¹ Just as Young 1998, 16, Bicchieri 2006, 15, maintains that social norms are equilibria. This implies that a normative rule is not a social norm if so many people violate it that it ceases to be

5 Conclusion

Social practices and institutions are interdependent behavioral regularities. According to the equilibrium approach, they are nothing more. It characterizes them as equilibria in coordination or cooperation games. I have criticized this approach, because it fails to satisfactorily explain how people coordinate their behavior or why they cooperate. I have argued that people coordinate their behaviors by means of signaling rules, and that they cooperate because of normative rules. In light of this, I have explicated social practices and institutions not only in terms of equilibria but also in terms of rules.

According to the Rules-*in*-Equilibrium theory (*RiE*), social practices are behavioral regularities that solve coordination problems by means of signaling rules. Institutions are social norms that govern social practices. According to the Rules-*and*-Equilibrium theory (*RaE*), they involve normative rules that stabilize patterns of coordination or enable cooperation, at least in principle. What the hybrid approach that I defend here illuminates is how the idea that social forms involve equilibria elucidates the roles social norms play, and the function that social practices and institutions have.

In order to give more substance to the idea that institutions are norm-governed social practices, I have presented an account of social norms and an account of what it means for them to govern practices. A social norm exists when people believe that they are supposed to conform to its normative rule. Although sanctions can motivate people to conform, the norm as such can do so as well. In order for it to do so, people have to regard the norm as legitimate. This in turn means that they take the beliefs and expectations on which the norm depends to be justified. Against this background, I have proposed that a social norm governs a social practice exactly if people are motivated to conform to it to a nontrivial degree, irrespective of whether this is due to sanctions or the norm itself.

Acknowledgment: I thank Francesco Guala, Chiara Lisciandra and Nic Southwood for insightful discussions and comments. I also acknowledge useful comments from one editor, Anton Leist, and an anonymous referee.

an equilibrium. Perhaps a behavior can indeed be the norm only if sufficient people conform to it. However, as I have just argued, the norm of an institution need not be in equilibrium.

References

- Aumann, R. J. (1974), Subjectivity and Correlation in Randomized Strategies, in: *Journal of Mathematical Economics* 1, 67–96
- Bandura, A. (2016), *Moral Disengagement*, New York
- Bicchieri, C. (2006), *The Grammar of Society*, Cambridge
- (2016), *Norms in the Wild*, New York
- Brennan, G./L. Eriksson/R. E. Goodin/N. Southwood (2013), *Explaining Norms*, Oxford
- Crawford, S. E. S./E. Ostrom (1995), A Grammar of Institutions, in: *American Political Science Review* 89, 582–600
- Gilbert, M. (1989), *On Social Facts*, London
- Gintis, H. (2007), *The Bounds of Reason*, Princeton
- Guala, F./F. Hindriks (2015), A Unified Social Ontology, in: *Philosophical Quarterly* 65, 177–201
- /L. Mittone (2010), How History and Convention Create Norms: An Experimental Study, in: *Journal of Economic Psychology* 31, 749–756
- Hindriks, F. (2009), Constitutive Rules, Language, and Ontology, in: *Erkenntnis* 71, 253–275
- (2013a), The Location Problem in Social Ontology, in: *Synthese* 190, 413–437
- (2013b), Collective Acceptance and the Is-Ought Argument, in: *Ethical Theory and Moral Practice* 16, 465–480
- /F. Guala (2015), Institutions, Rules and Equilibria: A Unified Theory, in: *Journal of Institutional Economics* 11, 459–480
- /— (forthcoming), The Functions of Institutions: Etiology and Teleology, in: *Synthese*
- Knight, J. (1992), *Institutions and Social Conflict*, Cambridge
- Lewis, D. K. (1969), *Convention: A Philosophical Study*, Cambridge/MA
- Prentice, D. A./D. T. Miller (1993), Pluralistic Ignorance and Alcohol Use on Campus: Some Consequences of Misperceiving the Social Norm, in: *Journal of Personality and Social Psychology* 64, 243–256
- Schotter, A. (1981), *The Economic Theory of Social Institutions*, Oxford
- Searle, J. R. (1995), *The Construction of Social Reality*, New York
- (2010), *Making the Social World. The Structure of Human Civilization*, Oxford
- Southwood, N. (2018), Laws as Conventional Norms, in: *Dimensions of Normativity*, ed. D. Plunkett/S. Shapiro/K. Toh, Oxford, 23–44
- Sugden, R. (1986), *The Economics of Rights, Co-operation and Welfare*, Oxford
- Tuomela, R. (2013), *Social Ontology: Collective Intentionality and Group Agents*, Oxford
- Ullmann-Margalit, E. (1977), *The Emergence of Norms*, Oxford
- Vanderschraaf, P. (1995), Convention as Correlated Equilibrium, in: *Erkenntnis* 42, 65–87
- Young, P. H. (1998), *Individual Strategy and Social Structure*, Princeton