

Status functions and institutional facts: reply to Hindriks and Guala

JOHN R. SEARLE*

Department of Philosophy, University of California, Berkeley, CA, USA

Abstract. Hindriks and Guala (2014) hope to provide a unified account of institutional theory that will combine the accounts of regulative rules, constitutive rules, and equilibria. I argue that only the constitutive rule approach has any possibility of success, and that the other two cannot even pose the right questions, much less answer them. Hindriks and Guala think constitutive rules can be reduced to regulative rules. I argue that their reduction is mistaken. The key to understanding social ontology is understanding status functions.

In their article, ‘Institutions, Rules, and Equilibria: A Unified Theory’, Hindriks and Guala (2014) claim they have identified three different approaches to institutions – institutions as game-theoretic equilibria, institutions as regulative rules, and institutions as constitutive rules. They hope to synthesize these into a single general account, and they wish to show that constitutive rules can be reduced to regulative rules. I will argue that their attempt to reduce constitutive rules to regulative rules fails, and that of the three theories only the third even begins to offer a possible account of institutions. The other two do not have an apparatus rich enough to answer, or even state, the questions. The problem with the first two is not just that they give the wrong answers to the questions, but that they cannot even pose the questions.

What then exactly are the questions? In human social life, we are perpetually immersed in institutional facts. I am a citizen of the US, this piece of paper is a twenty dollar bill, the car in the driveway is my property, I promised the editor of this journal to produce an article, and I just watched a football game on television, to take five examples. Institutional facts do not just involve money, property, citizenship, promises, and games, but government, marriage, universities, cocktail parties, summer vacations, sabbatical leaves, voting rights, income tax, grocery stores, and a thousand other institutional structures. What do all of these have in common that makes them institutional? *All institutional facts are status functions, and all institutions are systems that enable the creation and maintenance of status functions.* A theory of institutions is a theory of status functions. Humans have the capacity to impose functions on objects

*Email: searle@cogsci.berkeley.edu

or people, where the function cannot be performed solely in virtue of the physical structure of the object or person in question, but require a collectively accepted assigned status, and the function can only be performed in virtue of that collective acceptance. My knife performs its function in virtue of its physical structure, but the twenty dollar bill does not perform the function in virtue of its physical structure, but in virtue of its collectively accepted status as money. Such phenomena I call *status functions*. And why are status functions important? Status functions are the distinguishing feature of human societies, and hence the essential distinguishing trait of human civilization. Animals have pair-bonding, power structures with alpha females and alpha males, but they do not have marriage, government, private property, or money. Status functions are power relations, and the powers are, as far as we know, peculiar to humans: they are rights, duties, obligations, permissions, authorizations, etc. I call these *deontic powers*. All institutional facts are status functions, and all status functions are deontic powers. And again, why are these so important? All deontic powers lock into human rationality in a way that gives people reasons for action that are independent of their desires. Once again this capacity, as far as we know, is peculiar to humans. No other animal can act on its obligations and against its desires. How are we to explain institutions in a way that explains all of this? The analysis requires at least the following:

- (1) Any theory of institutions has to begin with a distinction between the *institution* as a *system* and *institutional facts* within the institution, between the institution of private property, for example, and the fact that this car is my private property. Because all institutional facts are status functions, we need to explain what an institution is and how it enables the creation of and maintains the existence of status functions.
- (2) How do institutional facts relate to brute facts? How, for example, does the institutional fact that this is a twenty dollar bill relate to the brute fact that it is a piece of paper?
- (3) What is the role of language in creating, constituting, and maintaining institutional facts? Because such facts only exist insofar as they are represented as existing, any theory of institutions has to contain a systematic account of representations, linguistic and otherwise, and their role in institutional reality.
- (4) What about language itself? It is certainly an institution with status functions. How does it resemble and differ from other institutions?
- (5) All institutional facts without exception create deontic powers: rights, duties, obligations, authorizations, permissions, etc. What exactly are these, and how do they function? This is the key feature of human institutions. One of the main differences between humans and other social animals is that humans can create and act on obligations and other deontic powers. Other social animals as far as we know cannot. How does it work?
- (6) On an intuitive account of deontic powers, they are the glue that holds human civilization together because they provide us with reasons for action that are independent of our desires. I have a reason to keep my promises even in cases

where I don't want to do the thing I promised to do and it is not in my interest to do it. How is that possible?

- (7) Specifically, if all voluntary actions are expressions of a desire to perform that action, how can desire-independent reasons for action be effective?
- (8) As far as we know, non-human animals have nothing like this apparatus. Why not?

These are some, but not all, of the questions. Any theory of institutions has to recognize and answer all of these and other related questions. I have attempted to answer all of them in a number of writings (Searle, 1995, 2001, 2005, 2010). As presented by Hindriks and Guala, the other theories do not even address them in a way that would enable us to get an adequate theory of status functions. Neither the regulative rule approach nor the equilibrium approach can even pose the questions, much less provide answers. One central set of questions is about the deonticity of status functions, and as far as I can tell neither the regulative rule theorists nor the equilibria theorists even addresses these questions.

There is one substantive claim that Hindriks and Guala make that has to be answered directly. They think that constitutive rules, rules of the form 'X counts as Y in context C', can be reduced to regulative rules, rules of the form 'Do X', and they attempt to carry out such a reduction in the case of the institution of property. I do not believe their account succeeds, and I want to briefly state why.

In their simplified case of what they call 'property*', where we have X counts as Y in C, then in an imagined community the first person to occupy the property has property rights, where 'occupy' is part of the X term and 'property' is part of the Y term, and spelling out the rights will consist of stating the rights that attach to that property. They are correct in seeing that there are two levels in the structure of institutions. There is a level of the X term and a level of the Y term. It can be a brute fact that I occupy the piece of land first, but that brute fact by the constitutive rule, we suppose, generates an institutional fact: it is my property. And that gives me property *rights*. How can that be reduced to a set of regulative rules? Here is their analysis. Hindriks and Guala (2014) say the constitutive rule that occupations count as creating property rights reduces to the following: There are two players, P1 and P2. 'The correlated equilibrium in the game of property* is the pair of strategies:

- (s1) Use if P1 occupied first, do not use if P2 occupied first.
- (s2) Use if P2 occupied first, do not use if P1 occupied first'.

This is stated using the imperative mood of the verb 'use'. The problems with this account are: first, that there is no such imperative in the actual human institution of private property, and second, there could not be any such an imperative even in the fakey artificial version of 'property*' that they have invented, because the very essence of property is a matter of *rights*, i.e. of a specific deontology. The institution of private property, even in the simplified version that they give us, would give the property owner *rights*, i.e. *deontic*

powers, where the owner is not under orders to do anything. The point is that their ‘correlated equilibrium’ ‘strategies’ do not just happen to fail to describe the institution of private property, but they could not describe any institution that was remotely like private property, and Hindriks and Guala (2014) implicitly concede this when they attempt to summarize (s1) and (s2) as:

‘[R] If one is the first to occupy a piece of land, one has the *right* to its exclusive use [*italics mine*]’.

One question is, how did that ‘right’ get in there? The two sentences, (s1) and (s2), in the imperative mood do not entail the existence of a right, and there is nothing to justify the claim that there is a right as stated in [R]. And such rights, as well as other deontic powers, are crucial features of institutional reality.

We know in advance that no system of regulative rules can be equivalent to constitutive rules, because the regulative rules always operate at the same level as the phenomena they regulate. By definition, regulative rules regulate antecedently existing forms of behavior. Drive on the right-hand side of the road, and eat with a fork in your right hand are examples. But constitutive rules introduce a new ontological level because the operation of a constitutive rule is invariably a speech act of declaration that a new institutional fact exists. It creates a new fact. It enables the user to count something at the lower level, the X term, as having the Y status function. In order to carry out their enterprise of reducing the constitutive to the regulative, you would have to show how a set of regulative rules at the same level generates a deontology of a higher level. It is impossible to do that and it is not surprising that they do not succeed or even make a serious effort. They see correctly that ‘X counts as Y’ gives two levels, the X level and the Y level, but what they do not see is that the X level by itself is not sufficient to generate the higher level, the Y level. The constitutive rule creates a new level of reality, a level of status functions with their deontic powers. They smuggle the deontology into their ‘summary’ of (s1) and (s2), but then the summary is no longer a matter of regulative rules. [R] is simply not a summary of (s1) and (s2), because it makes a reference to rights, which do not exist in (s1) and (s2). There has to be some procedure for *counting* a piece of land as my property with the accompanying deontology that goes with the ownership of private property.

It is also worth pointing out that the actual ‘rule’ they propose is out of the question as an explanation of property rights. Anyone reading this article has lived in systems of private property all their lives. Have they ever lived in a community where property was allocated by collective intentionality on the principle: the first person who occupies it owns it? The closest we came in American history was the Homestead Act of 1862 whereby if you occupied, lived on, built a house, and farmed a quarter section of land¹ for five years, you

1 ‘Section’ is a technical term. A quarter section is an officially subdivided unit of land of 160 acres.

owned it. But even in this case the principle was created, imposed, and enforced by federal law. There was no Nash equilibrium. It was a federally imposed distribution of property, and as one would expect, people cheated like crazy. One form of cheating was that several members of the same family would apply for adjacent quarter sections. Instead of small family farms, as intended by the law, they had one large family farm. Prior to the existence of federal law, the whites simply stole property from Native Americans. Is that supposed to be a Nash equilibrium?

The rule that they propose seems more appropriate for getting a seat on a bus. Whoever gets there first and sits down has the seat. But rather more is involved in the distribution of territory.

So their account of property rights is inadequate. What about the regulative rules and the equilibria accounts in general? As presented by Hindriks and Guala, each is hopeless as an account of institutions because they do not give us a rich enough apparatus to generate institutional facts. I believe Hindriks and Guala implicitly concede that regulative rules by themselves will not do the job unless constitutive rules can be reduced to regulative rules. But if I am right that no such reduction is possible, then the regulative rules account fails for Hindriks and Guala. What about games theory? It is out of the question to suppose that game-theoretic equilibria are either necessary or sufficient for institutions. In Hume's famous example of men rowing a boat, it is, I guess, a game-theoretic equilibrium for them to coordinate their efforts, but there is no institution that they have created of rowing a boat. It ought to arouse our suspicions that the favorite examples of institutions that are game-theoretic equilibria are not institutions at all. The favorite example is driving on the right in the U.S. or on the left in Britain. (As the theorists love to say, there are two equilibria.) It does not matter whether we drive on right or the left as long as everyone does the same thing. But in the sense in which government, money, marriage, and private property are institutions, driving on the right is not an institution. Why not? It is a legally required practice in the US, but it does not generate deontologies in the way of institutions like private property. It does not matter whether we use the word 'institution', but it does matter whether we see the differences between the constitutive rules of property and the regulative rule of driving on the right.

It is incidentally another flaw in their account that they treat normativity and deonticity as equivalent. They are not. All rules are normative. But not all rules create, for example, obligations.

Furthermore, there are lots of ongoing institutions for which there are no equilibria. Or if there is, it is totally irrelevant to the existence of the institution. Consider again the international institution of private property. Everywhere I go in the world, rights of private property exist. Is this because there is a game-theoretic equilibrium? In their toy examples of Nuer and Dinka, it is in everybody's advantage to grant the rights of the other. But what happens when it is no longer to mutual advantage? In real life, there are all sorts of disadvantages

to other people having so much property. Massively unjust distributions are everywhere. I have no idea how one would go about settling the issue of a Nash equilibrium involving several billion people, but that is beside the point. The recognition of property rights already gives desire-independent reasons for action. No Nash equilibrium could exist nor is it necessary that it should exist.

Because the entire equilibria approach seems to hang on this issue, let us spell it out in detail. Ignoring legal enforcement for the moment, my only reason for driving on the right in the U.S. is that everybody else does it, and there is common knowledge that everyone drives on the right. Given these facts, no one can benefit by driving on the left. There is an equilibrium. But now try it for promising. The parallel has to be, 'My only reason for keeping my promises is that everyone else does it and there is common knowledge that everyone else does'. But this is false. Kant was right that if everybody broke all their promises all the time, the institution would cease to exist. But my obligation to keep my promises is independent of whether you keep yours. In my experience, a high percentage of people break their promises. And for all I know, most people break most of their promises most of the time. But that is irrelevant to my obligation to keep my promises. What is the source of the obligation? When I make a promise, I intentionally create an obligation, a desire-independent reason for action. I create a new institutional fact, the fact that a promise exists. Game-theoretic equilibria are simply irrelevant. Let me emphasize this point: *You cannot do an equilibrium analysis of institutional facts of the sort they propose, because the equilibria are insufficient to generate the deontology – rights, duties, obligations, etc. – that is the defining trait of institutional facts.*

By the way, the mistake of the equilibria theorists was made by Hume in the 18th century. He thought we could give such prudential accounts of promising and property rights. I had supposed his accounts were long ago recognized as refuted, but the same mistakes apparently are being repeated today.

I have never seen an adequate equilibrium account of property, promising, government, or marriage, because in the case of the ones I have seen, they have an incoherent conception of rationality. They cannot explain the operation of desire-independent reasons for action, because the authors think there cannot be any such things. Rationality can require that I perform actions that I do not want to perform, that are not in my interests, and yet I have to perform them because I am, for example, under an obligation. On the standard maximizing conception of rationality, such situations are impossible. The adherents have to pretend in such cases that promise keeping is a way of maximizing utility. The pleasure in promise keeping is no different from the pleasure in drinking beer. It is just less fattening.

It is a non-trivial question how obligations can motivate, but they clearly do, and I have tried to explain how in *Rationality in Action* (Searle, 2001). This is not the place to summarize the whole argument, but the basic principle is that

the recognition of the validity of a desire-independent reason can be the basis for forming a desire to perform the act. In the case of beer drinking, the desire is the basis of the reason, indeed it is the reason. In the case of obligations, the reason is the basis of the desire. But of course this does not always work: lots of people recognize the validity of their obligations but are unmoved to act. The most obvious cases where the principles do work are epistemic: there is something I very much do not want to accept or believe, but the evidence is overwhelming, so I have a desire-independent reason for forming a desire to accept something I do not want to accept.²

As far as I can tell from Hindriks and Guala, the contribution of equilibrium accounts to the theory of institutional ontology is zero, nothing; and this is because they are addressed to a different question, the question of how we can show that certain sorts of coordinated behavior satisfy Nash equilibria.

Typical game-theoretic accounts presuppose the existence and nature of institutions and then try to give grounds for selecting an institution. They assume we know what property is and then offer, as Hindriks and Guala do, ‘strategies’ for ‘correlated equilibrium’. But the question that has got to be answered is: What is property in the first place? The endless discussions of the ‘prisoners’ dilemma’ are examples of this built-in superficiality. The typical discussion begins: ‘A prison warden offers two prisoners the following choice . . .’ But what is a prison? What is a warden? What is a prisoner? What is an offer? These questions are never asked. You get a totally different conception of the prisoners’ dilemma if you think of it as the ‘lovers’ dilemma’, ‘the mother and child dilemma’, ‘the Mafiosi’s dilemma’, and so on. If future generations of intellectuals find our sight endurable at all – not, I admit, a likely prospect – they will be flabbergasted by the amount of attention paid to the prisoners’ dilemma.

A remarkable thing about human institutions is that they are typically built on top of a reality of brute facts, but they only exist insofar as they are represented. This means that a specific theory of language is essential for a theory of institutional facts, and not any representation will do. Hindriks and Guala see correctly that human institutional facts only exist insofar as they are represented as existing, so they add to their account a proviso that there must be representations. This, they think, will enable them to eliminate possible counterexamples from animals that have game-theoretic solutions to coordination problems but do not have institutions. But that is not good enough. They need to provide a theory of representations that will explain their functioning institutional ontology. Are representations institutional facts? And what sort of an institution is the institution of representing?

2 Kant in the *Groundwork of the Metaphysics of Morals* (1964, pp. 122–123) faces the same problem, but his answer is disappointing. He says we cannot explain how the categorical imperative motivates, but it does motivate by creating a feeling of pleasure. How does this work? He cannot say, because it involves the operation of the world of things in themselves. For him it remains a mystery.

As presented by Hindriks and Guala, neither the equilibria account nor the regulative rules account even gets started as a general attempt to answer the questions I listed above. Why would anyone think otherwise? Why would anyone think regulative rules or equilibria provide the essential ontology of institutional reality? I do not know enough of the literature to answer this question, but there appear to be a number of reasons. First, the authors are not much interested in social ontology. They are interested in economics or in games theory, but not in the ontology of human social reality. Second they have an inadequate conception of rationality. It was the theory taught in Economics in my childhood. Long refuted, but still around. An unfortunate feature of the Hindriks–Guala account is they make it look as if there are only three candidates for a theory of institutional ontology. But if I am right, two of those are non-starters. Hindriks and Guala leave out a whole lot of other powerful theories. They mention some, but not all, in passing in their footnotes, but the theories in question deserve more attention. Discussion of these issues ought to take seriously the work of Margaret Gilbert, Geoffrey Hodgson, Tony Lawson, and Raimo Tuomela among others. They at least recognize the problems of social ontology.

References

- Hindriks, F. and F. Guala (2014), 'Institutions, Rules, and Equilibria: A Unified Theory', *Journal of Institutional Economics*, published online: 16 October 2014, DOI: [10.1017/S1744137414000496](https://doi.org/10.1017/S1744137414000496), 1–22.
- Kant, I. (1964), *Groundwork of the Metaphysics of Morals*, New York: Harper & Row.
- Searle, J. R. (1995), *The Construction of Social Reality*, New York: Simon and Schuster.
- Searle, J. R. (2001), *Rationality in Action*, Cambridge MA: MIT Press.
- Searle, J. R. (2005), 'What is an Institution?', *Journal of Institutional Economics*, 1(1), 1–22.
- Searle, J. R. (2010), *Making the Social World*, Oxford: Oxford University Press.