Article

Author's Copy

Understanding Institutions without Collective Acceptance?

⑤SAGE

Pekka Mäkelä[1], Raul Hakli[1], and S. M. Amadae[1]

Abstract

Francesco Guala has written an important book proposing a new account of social institutions and criticizing existing ones. We focus on Guala's critique of collective acceptance theories of institutions, widely discussed in the literature of collective intentionality. Guala argues that at least some of the collective acceptance theories commit their proponents to antinaturalist methodology of social science. What is at stake here is what kind of philosophizing is relevant for the social sciences. We argue that a Searlean version of collective acceptance theory can be defended against Guala's critique and question the sufficiency of Guala's account of the ontology of the social world.

1. Introduction

In his impressive book, *Understanding Institutions: The Science and Philosophy of Living Together*, Francesco Guala proposes a new unified theory of social institutions. The argumentation in the book is admirably versatile and covers a wide range of topics both in social sciences and in philosophy, from philosophy of social sciences to metaphysics. Guala avails himself of

Corresponding Author:
S. M. Amadae, University of Helsinki, Helsinki, Finland.
Email: sm.amadae@helsinki.fi

conceptual and theoretical tools of equilibria accounts and rule-based accounts of social institutions, and with these tools at hand, he constructs his own unified account of social institutions.

Even if the main topic of the book is the nature of social institutions, the book importantly contributes to more general issues and debates in the philosophy and methodology of social sciences. As Guala unfolds his account, he ends up defending a view of the ontological nature of the social reality, which makes the social world an apt object of scientific study. The criteria of proper objects of scientific study are given in terms of natural and real kinds grounding projectability. The positive account provided in the book would be a respectable achievement in its own right but Guala does not content himself with this but he also provides detailed and lengthy critiques of alternative and competitive approaches and accounts. In particular, he challenges the defenders of the collective acceptance views, such as Searle (1995, 2010) and Tuomela (2007, 2013), by way of arguing that their view leads to antirealism and infallibilism about social institutions and institutional facts.

As a consequence of the multifaceted aims and results of Guala's book, the stakes are set relatively high for the ensuing debate and argumentation. One of the central issues is whether social reality can be an apt object of scientific study. In Guala's view, the collective acceptance accounts of social institutions are committed to ontological or constitutive dependence on representations, which renders them objects such that (a) knowledge of them can be attained directly, without risk or error, and that (b) the knowledge of them cannot be used to make inferences like those in natural sciences. Hence, they entail antirealism and infallibilism about institutions. If this were true a large portion of literature in the field of social ontology and collective intentionality (that studies social phenomena including social institutions, collective action, and collective attitudes such as collective beliefs, collective intentions, etc.) would be futile from the perspective of serious social-scientific research, these views would commit their proponents to antinaturalist or separatist view of the methodology of social sciences.

In what follows, we, first, lay bare the core elements of Guala's argumentation against the acceptance views of social institutions; second, we discuss argumentative strategies available for acceptance view theorists to defend their approach; third, we study in detail what we take to be the most viable lines of argumentation to save the acceptance views from Guala's critique; fourth, we discuss some tentative critical points concerning the capacity of Guala's unified theory to fully account for all the features of institutional reality considered central by acceptance view theorists; and fifth, we conclude that Guala's contribution in *Understanding Institutions* significantly advances the current discussion in many respects but is not yet the last word on the nature of social institutions.

2. Guala's Account and Critique

The multidisciplinary literature on institutions has a feature that makes strict and direct evaluation of the argumentation rather difficult, that is, it has a moving target as it were. Different authors have slightly different classes of phenomena in mind when they talk about institutions.

As an example, Guala and Hindriks (2015) express a very broad view of institutional reality:

> The social world is populated by entities such as norms, conventions, customs, laws, organizations, groups, identities, and roles. Like other theorists, we will refer to these seemingly diverse things using the generic term "institution" together with its variants, like "institutional role" and "institutional fact." (Guala and Hindriks 2015, 177)

In contrast, Searle (1995, 2010) has a narrower conception of institutional reality: Searle's (1995) focus is on institutional facts, such as "this piece of paper is a twenty-dollar bill," and more general classes of things such as money, property, marriages, and governments. Arguably, such things have a function that is not performed in virtue of sheer physics but in virtue of collective intentionality (Searle 1995, 39). Social institutions are constructed by human beings (some group) by way of collectively accepting (for themselves) that some things in the world are seen as or taken to be something they are not by their physical character. This presupposes collective intentionality, status functions, deontic powers, and so on. Hence, institutional reality is a subclass of social reality. Similarly, in 2010 he writes,

> The distinctive feature of human social reality, the way in which it differs from other forms of animal reality known to me, is that humans have the capacity to impose functions on objects and people where the objects and the people cannot perform the functions solely in virtue of their physical structure. The performance of the function requires that there be a collective recognized status that the person or object has, and it is only in virtue of that status that the person or object can perform the function in question. Examples are pretty much everywhere: a piece of private property, the president of the United States, a twenty-dollar bill, and a professor in a university are all people or objects that are able to perform certain functions in virtue of the fact that they have collectively recognized status that enables them to perform those functions in a way they could not do without the collective recognition of the status. (Searle 2010, 7)

In Guala's usage "institution" appears to refer to a broader class of phenomena belonging to social reality than in Searle's. For instance, Guala (2016) includes conventions such as driving on the left-hand side to the class of institutions, whereas Searle does not, because they lack the features that he

considers necessary for institutions, in particular, functions that cannot be performed solely in virtue of their physical properties. Obviously, this makes it a bit difficult to evaluate which one is offering a better and more adequate account. However, this problem can be overcome: if Guala's theory can account for all the aspects Searle aims at accounting for, then Guala's broader scope does not seem to be a fundamentally serious problem. So this is not a linguistic point but about the ontology of the elements of social reality.

One question we have been pondering is whether Guala's account can satisfactorily explain the ontological step that understanding of the institutional reality arguably requires. Let us think about a very simple example of a traffic light, which is a "conventional" coordination device. Before going to the example itself, it is perhaps appropriate to point out that it is not supposed to be an example of social entities or objects that are of central interest to accounts of social ontology and collective intentionality. Rather it is meant to illustrate a central mechanism in the construction of social reality that involves collective acceptance and is required even in simple cases of coordination devices. For a traffic light to perform and satisfy the function of a coordination device, it has to be interpreted in a certain way, namely, as having certain deontic meanings (expressing rights or obligations) attached to colors. Obviously, red does not convey the message "stop" just in virtue of its physical features, this message is of human making. The question concerns the nature of the "human touch" here. In light of Searle's account, a collective imposition of a status function has taken place, members of the population have collectively accepted that the color red in a traffic light means "stop," they have given a new status with deontic powers to the color. In other words, they have a constitutive rule as a content of their collective acceptance, and this collective acceptance presupposes collective intentionality. Guala, in contrast, would appeal to an equilibrium account for the phenomenon. According to him, the traffic light can perform the function of a coordination device perhaps for its salience or tradition, and this is a way to understand what a traffic light is. However, this move seems to sidestep the ontological question explicit in Searle's account. Beliefs and expectations about other people's behavior make agents behave in a certain way: when these are collectively aligned in a certain way, say in a coordination game kind of situation, it so happens that beliefs and expectations of certain kinds of behavior "cause" that kind of behavior and the beliefs and expectations also become true. Behavior and strategies can be in an equilibrium, and it can be an equilibrium to treat certain kinds of objects as having properties that they do not have as physical entities. If reaching an equilibrium presupposes a coordination device and such a device presupposes a status function imposition, then we might say that Guala's account presupposes Searle's account.

Guala and collective acceptance theorists seem to agree upon quite a lot; however, there is a question whether Guala has an account of objects having properties not based on their physical properties, which seems to be the fundamental ontological move in Searle's theory (or an argument that no such things are needed).

In what follows, we will mainly be interested in two different issues: one is whether and how seriously Guala's critique undermines the tenability of collective acceptance view theories, the other is whether the positive account provided by Guala is successful in the sense of being able to explain all the relevant features of institutional reality.

Guala has a wide selection of critical ammunition against collective acceptance theories, and we will not touch upon all the critiques presented in the book. In what follows, we focus on one line of critical argumentation, which we find to be well-targeted and to have a potentially devastating impact, namely, Guala's argument that collective acceptance views lead to antirealism and infallibilism concerning social institutions. The reasoning in brief runs as follows: (a) collective acceptance views of institutions and institutional facts are committed to a dependence thesis (to be discussed below); (b) the commitment to the dependence thesis binds the acceptance views to antirealism and infallibilism; (c) in virtue of this boundedness to antirealism and infallibilism the understanding of institutions and institutional facts provided by collective acceptance views excludes institutions from the class of apt objects for social-scientific study. Guala, however, argues that the dependence thesis, the commitment of collective acceptance views, is not needed to account for social institutions and institutional facts, and he thus saves these phenomena as proper objects for social-scientific study.

As to the background of Guala's argumentation, he points out that when we talk about real kinds, we mean either that they exist independently of our theories, beliefs, representations, or that they are genuine kinds, in the sense that they support inductive inferences and generalizations (Guala 2016, 146). Guala argues that these two senses of "real" are related: in order for a kind to be a proper object of scientific study, it has to support inductive inferences and generalizations, and that requires that the kind does not depend, in a problematic way, on our theories, beliefs, or representations. According to Guala, there are two types of mind dependence. One, namely, causal dependence, does not undermine real kinds and is therefore metaphysically innocent. For example, interactive kinds are such that they depend on representation, but the dependence is causal: our classifications of people may change the ways in which these people experience themselves, which may lead them to change their behavior (Hacking 1999, 104). As a result, the interactive kind will change, but this change does not make the kind unprojectable.

However, Guala claims that constitutive (or ontological) dependence is problematic. For instance, when Searle says things like "money is money because it is believed to be money," he seems to be committed to a view that something is true in virtue of being believed to be true. This inspires the anguish that these views are in the idealistic camp. More generally, according to Guala, the collective acceptance views endorse the dependence thesis of the form: necessarily, X is $K \rightarrow CA(X$ is $K)$

This endorsement, according to Guala (2016, 150), is a problematic move because arguably it is bound to lead to a violation of either the principle of metaphysical innocence or the principle of realism. Boyd presents the principle of metaphysical innocence as follows:

> . . . human social practices, like the adoption of theories and classificatory schemes, are metaphysically innocent: they affect the causal structure of the world only via the operation of intermediary causal mechanisms which supervene on the causal structures studied by the various special sciences and not also in some additional way studied only by philosophers practicing conceptual analysis. (Boyd 1991, 144-45)

As to the principles of realism, Guala (2016, 151) draws on Thomasson's (2003, 583) formulation according to which they are the following:

> **Extensionality:** there is a kind with natural boundaries that determine the extension of the term independently of anyone's concept(s) regarding the kind.

> **Error principle:** since these boundaries are not determined by human beliefs about those boundaries, any beliefs (or principles accepted) regarding the nature of Ks could turn out to be massively wrong.

> **Ignorance principle:** for all conditions determining the nature of the kind K, it is possible that these remain unknown to everyone.

The obvious question is why would the violation of these principles follow from the endorsement of the acceptance thesis. Let us start with the violation of one of the principles of realism, namely, extensionality. The reasoning seems to run along the following lines: the collective acceptance thesis entails ontological dependence, X depends ontologically on Y equals "necessarily $X \rightarrow Y$." The collective acceptance thesis appears to fall under this structure, because, according to it, X's being K ontologically depends on it being collectively accepted that X is K. From this, it seems to follow that the boundaries of Xs that are K are not determined independently of any

representations regarding X's being K, indeed, it seems that the collective acceptance thesis explicitly states that these boundaries are determined by the collective's representation.

What makes such violation of the extensionality principle harmful and harmful from which perspective? Perhaps the idea is that the violation is harmful because it implies infallibilism, to put it bluntly, because X's being K depends on the collective believing that to be the case, the collective cannot be wrong about it. One would think that such infallibilism is worrisome from the perspective of scientific study, as infallibilism does not sit nicely with falsifiability. Hypotheses about the nature of money could not be subjected to empirical tests, rather their truth could be verified by probing the contents of our own beliefs.

Who endorses the collective acceptance thesis? According to Guala, the thesis is endorsed by collective acceptance theorists. Who are the ones riddled with infallibilism? They are the members of a community, who, according to collective acceptance theories, create the institution for themselves (the community having money cannot be massively wrong about their having money). In our view, nothing of what has been said so far makes any science studying institutions infallible about institutions.

The basic idea in Guala's argument is that if we endorse the accounts of social institutions that are based on collective acceptance, we are committed to antirealism and infallibilism with respect to institutional kinds. That is, we are forced to accept the following argument:

1. Real kinds are ontologically independent of our representations
2. Science can only study real kinds
3. Institutional kinds depend ontologically on representations
4. Hence, institutional kinds are not real kinds (from 3, 1)
5. Hence, institutional kinds cannot be studied by science (from 2, 4)

But, according to Guala, institutional kinds can in fact be studied by science, so we should deny collective acceptance theories of social institutions. Our aim is to show that this argument fails, and we have several alternative strategies to show that. Our first strategy denies premiss 3. The second strategy denies 1. The third strategy denies premiss 2. We will go through them in the sections that follow.

3. Ways to Block Guala's Argument

In this section, we will discuss possible options for collective acceptance theorists to resist Guala's argument:

0. Deny the validity of the metaphysical innocence principle.
1. Deny the dependence thesis as stated by Guala.
2. Accept the dependence thesis but argue that it does not entail antirealism and infallibilism.
3. Accept the dependence thesis and that it entails antirealism and infallibilism but argue that this does not prevent institutional kinds from being proper objects of social-scientific study.
4. Accept that institutional facts and institutions may not be proper objects of scientific study but claim that social ontology is a respectable project of philosophical study along the lines of Strawsonian descriptive metaphysics (unfolding the features and commitments of our conceptual framework).

In the following subsections, 3.1 and 3.2, we will focus on options 1 and 2, which we find most promising. As to the principle of metaphysical innocence (option 0), Guala mentions that he is not aware of anyone who would have made a serious argumentative attack against this principle. We do not find an attempt to deny the validity of the metaphysical principle a viable strategy for a collective acceptance theorist either, rather the question worth discussing is whether all the sensible readings of collective acceptance theories are bound to violate the principle. We are inclined to claim that this is not the case. It is not totally clear to us what Guala and Boyd have in mind in invoking the principle of metaphysical innocence. We are not familiar with a philosopher engaged in conceptual analysis who believes they are affecting the causal structure of the world just by way of doing conceptual analysis. Clearly at least Searle and Tuomela have tried to tell a causal story about how the collectively accepted institutional reality affects people's behavior, namely, by creating desire-independent reasons for action which, once recognized as such, influence people's decision-making. Options 3 and 4 will be briefly discussed in subsection 3.3.

3.1. How to Formulate the Dependence Thesis

Recapitulating: According to Guala collective acceptance theorists, for example, Tuomela and Searle, are committed to an ontological reading of the dependence thesis. The worry here is that ontologically dependent objects are unlikely to be projectable and cannot be used for scientific purposes. Guala formulates the dependence thesis attributed to collective acceptance theories as follows (Guala 2016, 163):

$$\text{necessarily, } X \text{ is } K \rightarrow CA(X \text{ is } K),$$

which seems to be equivalent to the following (with the box denoting necessity):

$$\Box\big(X \text{ is } K \rightarrow CA(X \text{ is } K)\big).$$

In our view, Searle's formula (X counts as Y in C) should instead be formalized as follows:

$$\forall i X(i) \rightarrow Y_C(i). \tag{1}$$

This is the content of what is collectively accepted, and hence the statement including collective acceptance is:

$$CA_C\big(\forall i X(i) \rightarrow Y_C(i)\big).$$

If this is correct, the claim that if there is something that is (or counts as) Y in C then it is collectively accepted as Y (in C) should be formulated as:

$$\forall i Y_C(i) \rightarrow CA_C\big(Y_C(i)\big).$$

In order to get there, one must presumably make the additional assumption that in order for something to be Y, it must have been collectively accepted as X and that collective acceptance is closed under modus ponens. These assumptions are needed in order to get from the collective acceptance of the implication sentence to the collective acceptance with only the consequent.

In order to translate the above formulas into the symbolism used by Guala, we will have to change the predicate X to a variable and replace the domain condition C (which in Guala's version means the condition accepted to determine K-hood) with G (for group). The dependence thesis then becomes:

$$\Box\big(\forall x K_G(x) \rightarrow CA_G\big(K_G(x)\big)\big).$$

However, this does not seem to be about the kind "money" but rather about what Guala calls token institutions ("currencies" in the monetary case). In order for something to be a currency, or money in some particular society, the society members have to accept it as their currency, that is, money for them. The extension of the kind "money" would then consist of all the currencies, and a dependence thesis with respect to that general kind would go as follows:

$$\Box\big(\forall x K(x) \rightarrow \exists g CA_g\big(K_g(x)\big)\big).$$

This would mean that if something is an instance of the kind money, then there must be a society that has collectively accepted it as money in their

society. This is something that collective acceptance theorists might be committed to, but is this sufficient for Guala's dependence claim? There is plenty of reason for doubt. The main problem is that the K-predicates are different: On the left-hand side there is the general kind predicate K, but on the right-hand side, there is the society-relative predicate K_g. Hence, the three theses, Extensionality, Error principle, and Ignorance principle, do not seem to be obviously threatened. There could be a kind "money," that is K, independently of anyone's concept regarding *that kind*, because what is needed is only the society-members' society-relative concepts K_g, which may be rudimentary, mistaken, and far-removed from the proper kind "money" that social science tries to discover. So the principles of Extensionality, Error, and Ignorance, might all be retained.

One option for the collective acceptance theorists is hence to deny the dependence thesis as Guala formulates it and replace it with a weaker kind of dependence in which the kind depends on collective acceptance that does not concern the kind itself (e.g., all currencies in all possible communities) but only something that is part of what constitutes the kind (e.g., one currency in one particular community). (For an interesting discussion on types and tokens of institutions, see Aydinonat and Ylikoski 2018.) However, assuming that there are collective acceptance theorists who share Guala's interest in the existence of general kinds and think that they are dependent on collective acceptance concerning the kind itself, this might not be a satisfactory solution. Let us therefore consider another option that admits that there can be dependence on representations concerning the kind itself and see whether that leads to antirealism.

3.2. Whether the Dependence Thesis Leads to Antirealism

According to Guala, science can only study real kinds, which are ontologically independent of our representations. In the previous section, we considered the option that denies the ontological dependence of institutional kinds. In this section, we consider the option that admits the ontological dependence of institutional kinds, as characterized by Guala, but argues that Guala himself is committed to the existence of real kinds that are ontologically dependent on representations. Here we rely on Guala's claim that interactive kinds can be real kinds despite their causal dependence on representation and argue that this causal dependence satisfies also his characterization of ontological dependence. This opens up the possibility that institutional kinds are real in spite of being ontologically dependent on representations. This would be in line with what Uskali Mäki (2005) says about realism and dependence: "While other forms of realism may employ the notion of mind-independence,

scientific realism needs ideas of science-independence." This allows the possibility of scientific realism about institutional kinds such as money.

As to the nature of the dependence, Guala thinks that dependence on mental representation is harmful in the case of institutions but not in the case of interactive kinds, because in the case of the latter the dependence is causal and

> we can tell a plausible story about the way in which the creation and maintenance of behavioral regularities is dependent on human representations. But under the non-causal or ontological interpretation it is not clear how these correlations can be ensured by means of representation. (Guala 2016, 149-50)

It seems clear to us that collective acceptance theorists are not defending purely noncausal dependence, their story is mainly causal, the question is how harmful the noncausal element is? Guala seems to think that if there is a noncausal element involved, then a causal story is excluded. To us this kind of thinking seems to entail a commitment to the logical connection argument, which used to be a topic of discussion in philosophy of action but which most philosophers currently reject. Roughly, according to it, if there is a logical and hence necessary connection between two things, then there cannot be a causal connection between them, because causal connections are supposed to be contingent. The standard objection is that even though there is a conceptual connection between cause and effect (if there is an effect there must be a cause), it would be absurd to claim that there cannot be a causal connection. Although there is ontological dependence between the concepts, every instance of cause and effect is causal. The case of institutions seems similar to us: X to be represented in a certain way is a necessary condition of X to be treated in a certain way, indeed X counting as Y entails that X is represented as Y and as a result of such representation treated as Y, and all this must be collective in the case of social practices. However, in each and every instance of such a practice it is true that institutional entities are causally brought about by the actions of the members of the community, and in such a process a specific kind of representation appears to be a necessary part of a causal story.

We can ask whether there is a similar noncausal element in play also in the case of interactive kinds. This seems to be the case if we follow Guala in his characterization of ontological dependence:

$$X \text{ depends ontologically on } Y = \text{necessarily, } X \rightarrow Y.$$

Here X and Y are propositions, but it might be better to understand them as predicates. Then the dependence of something's being X on something's being Y could be written as

$$\left(\forall i X(i) \rightarrow \exists j Y(j)\right).$$

This says that whenever there is something that is X there must be something that is Y.

Guala has argued that this sort of dependence holds for institutional kinds, because for something to be an instance of an institutional kind K, there must be a group of people who collectively accept it as an instance of kind K, and hence have a representation of it as an instance of K. Guala notes that interactive kinds are in a way dependent on representation as well, but that dependence is causal instead of ontological. However, given the way Guala characterizes ontological dependence, such dependence seems to hold for many interactive kinds as well: in order for a classification of a person as a member of a kind to have an effect, someone—perhaps the person herself—must represent her as belonging to that kind. In Guala-like formalization, this means something as follows:

$$\text{necessarily, } X \text{ is } K \rightarrow Bel\left(X \text{ is } K\right),$$

where we take Bel to denote a belief of the person that must exist for the classification to have any effect. We can put this a bit more precisely as follows:

$$\left(\forall i K(i) \rightarrow \exists p Bel_p\left(K(i)\right)\right).$$

Hence, Guala's characterization seems to entail that those interactive kinds are ontologically dependent on representations concerning the kind. It would seem that institutional kinds and interactive kinds stand or fall together: If Guala wants to hold on to interactive kinds being real, he will have to admit that real kinds can be ontologically dependent on representations. Otherwise he would have to accept antirealism and infallibilism concerning interactive kinds.

However, Guala makes a different move in the article "On the nature of social kinds." There he says that not any kind of attitude dependence will suffice for realism to be challenged. Instead, "the attitude in the formula must be directed toward K, by stating conditions of kindhood, that is, by specifying the properties that make X belong to K" (Guala 2014, 61). The formula 2 below states this idea.

So instead of the case of a token entity accepted as a member of K,

$$X \text{ is } K \rightarrow CA\left(X \text{ is } K\right),$$

Guala (2016, 164) considers the case with a more general condition C accepted to determine whether something counts as K:

$$X \text{ is } K \rightarrow \left[CA\left(X \text{ is } K \text{ if } C\right) \& C \right]. \tag{2}$$

To return to the formalism used in the previous section, the implication of formula 2 above should be

$$\forall x K_G(x) \rightarrow CA_G\left(K_G(x)\right) \& C(x). \tag{3}$$

Or $\forall i Y_C(i) \rightarrow CA_C\left(Y_C(i)\right) \& X(i)$ with variable symbols closer to Searle's original formula.

Here it is important to note that Guala's use of C conditions is different from Searle's, the C conditions in Guala's sense are built into the X term in Searle's formulation. For Searle, C is basically just the society in question, not conditions of moneyhood. Moreover, conditions of moneyhood in Searle's formula are not, we think, meant to be conditions of the kind money, but rather just conditions for recognizing instances of money in the respective society. These conditions are such that, in the case of C being the United States, they set apart certain pieces of paper, namely, U.S. dollars, from other pieces like Euro notes: U.S. dollars $(= X)$ count as money $(= Y)$ in the United States $(= C)$. These conditions have little to do with the conditions that determine whether something is an instance of the scientific concept "money" that the social scientists are interested in. And these conditions are conventional. Conditions of moneyhood are instead involved in predicate Y so in order to participate in collective acceptance, individuals must have some understanding of what money is, but that understanding can be very rudimentary (e.g., that you can buy stuff with it) and the individuals may be to a large extent ignorant or mistaken about what makes something fall under the kind money in the social-scientific sense.

3.3. The Remaining Options

The remaining options for the collective acceptance theorists would be to either turn antirealist or Strawsonian. Amie Thomasson (2003, 606-07) takes the former option. She argues that despite antirealism and infallibilism, there is room for social science:

  1)"the kinds may be opaque to those not involved in the production or maintenance of those entities within the context of the relevant society."

2) "even within a community of insiders . . ., there are many facts to be uncovered by social science . . . e.g., the Marxist and feminist claims about (perhaps) unintended and unnoticed oppressive consequences of our practices involving money, division of labor, etc. Such causal facts certainly remain opaque and in need of discovery. There is even room for critique of elements of a society's metaphysical understanding of its own institutions, e.g., in exposing the beliefs of a society that believes that its institutions (kings, laws, customs) are established through natural or supernatural powers rather than simply through collective acceptance. The range of social scientific discovery remains as wide as we ever expected."

Alternatively, the project of Searle and Tuomela could be interpreted as an endeavor in Strawsonian descriptive metaphysics. A rough description of the Strawsonian view of descriptive metaphysics is as follows (see Glock 2012):

Strawson sets descriptive metaphysics the task not of limning the necessary structure of reality, but of elucidating "our conceptual scheme," "the way we think of the world," or "the actual structure of our thought about the world" (Strawson 1959, 15, 9). By contrast to Quine and the current naturalistic mainstream, Strawson insists that philosophy is not part of [or] continuous with science. In line with Wittgenstein, he maintains that philosophy should not try to rival science by describing or causally explain reality; instead it should elucidate our conceptual framework.

Here the aim of the ontological analysis is to make explicit the ontological commitments of the common sense view and folk psychology framework, it is transcendental argumentation for the view of the world to which we must be committed in light of our beliefs about the world we live in. If this would be correct then even if the only service these analyses can do for the social sciences is to vaguely point to social phenomena to be understood and explained in scientific terms it would not actually take much out of their philosophical interest or respectability. It would of course be bad news for those who think that social ontology deserves recognition only if it is in fruitful interaction with actual social sciences. Still, there may be some promise here.

4. Comparison of the Accounts

4.1. Points of Agreement

It is interesting to compare the following quotes, from Searle and Guala. It seems that there is no very stark contrast between them. Guala (2014, 65) writes,

what matters is not what type of attitude people have toward a certain class of entities (the conditions they *think* the entities ought to satisfy in order to belong to that class), but what they do with them in the course of social interaction. The relevant attitudes, in other words, are directed toward the attitudes of other people.

Later Guala (2014, 66) continues,

So to the extent that theoretical terms like "property," "money," or "professor" refer to something real, they refer to profiles of actions. The real content is not in the C conditions ("issued by Central Bank"): it is in the strategies ("accept it as payment") that are associated with the theoretical term ("money"). The kind money ultimately is nothing but this set of actions and related set of expectations. The C conditions are useful in so far as they simplify our decisions: they are *coordination devices* that help us identify quickly and without lengthy inspection an appropriate set of actions in the given circumstances. (Should I accept a piece of paper as payment? Yes, because it has been issued by the Central Bank.) But to focus on the C conditions as what makes something a member of K is a perceptual mistake. It mistakes the coordination device for the system of actions and expectations that a social institution is.

In comparison, Searle (1995, 36) writes,

It is tempting to think of social objects as independently existing entities on analogy with the objects studied by the natural sciences. It is tempting to think that a government or a dollar bill or a contract is an object or entity in the sense that a DNA molecule, a tectonic plate, or a planet is an object or an entity. In the case of social objects, however, the grammar of the noun phrases conceals from us the fact that, in such cases, process is prior to product. Social objects are always, in some sense we will need to explain, constituted by social acts; and, in a sense, the object is just the continuous possibility of the activity. A twenty dollar bill, for example, is a standing possibility of paying for something.

He also states (Searle 1995, 56-57):

The explanation for the apparent primacy of social acts over social objects is that the "objects" are really designed to serve agentive functions, and have little interest for us otherwise. What we think of as social objects, such as governments, money, and universities, are in fact just placeholders for patterns of activities. I hope it is clear that the whole operation of agentive functions and collective intentionality is a matter of ongoing activities and the creation of the possibility of more ongoing activities.

So both Guala and Searle emphasize people's activities. They are ultimately what make certain pieces of paper money. Guala (2014, 67) continues,

> The truly important properties—those that turn a token piece of paper into money, for example—are not conventional at all: they involve facts like people's beliefs about the likelihood that others will accept paper bills in exchange for goods and services.

In our view, these collectively shared beliefs *instantiate* or *exemplify* the *collective acceptance* which is *necessary for such pieces of paper to have a property or properties* they would not have on the basis of their physical properties, this is the *ontological step*, and this makes it possible for such entities like money to exist and to be objects of scientific study. A theory of money, stating true conditions of money, would not be possible without such street level ontological steps, nor would theories about other parts of institutional reality.

4.2. Points of Disagreement

What is the content of collective acceptance in Guala's interpretation and what is it in Searle's original account? For one, the condition C is understood differently in Searle's account that it is in Guala's account, and Guala (2014, footnote 8) is explicitly aware of this. Despite being aware of the different formulation of the "magic formula," it seems that Guala assumes that his formulation makes the intention of Searle's account more explicit and differs from it only in philosophically insignificant ways. However, C in Searle's formula roughly plays the role of relativizing the collective acceptance and the validity of its content to the accepting community, "for us" in Tuomela's lingo, whereas in Guala's account it at least seemingly has a different role. C in Guala's reading of Searle expresses the conditions that X needs to satisfy to count as Y, and as such it has a heavy semantic-conceptual role (conditions of Y-hood) and also its content, for example, in the case of money seems to presuppose an edifice of institutional structures. For instance, if C reads as something like "issued by the Central Bank" it already presupposes the whole monetary system of a state. We think it is obviously a contingent and conventional feature that for a piece of paper to count as money it has to be issued by the Central Bank, and so thinks Guala as well, and to our understanding so does Searle.

Indeed, it is not an easy task to be explicit about the content of the collective acceptance required for the existence of institutional facts in Searle's account. Searle discusses the content in terms of the magic formula which is

generic "X counts as Y" and when it comes to examples he contents himself with something vague like "this piece of paper counts as a twenty dollar note." However, this vagueness is understandable if we think that Searle's primary point is to emphasize that institutional facts require the collective imposition of status functions, that is, the collective practice of treating something as having a function and properties that it does not have on the basis of its physical nature. This according to Searle holds across the board of institutional facts. Searle does not seem to be worried about the exact propositional content of the collective acceptance.

Guala (2014, 62-64) argues that collective attitudes are not necessary for institutions. The first step in his argument is as follows: "I will argue that we can all be wrong about the nature of any institutional kind." This is easy to accept but it is a bit more difficult for us all to be wrong about the nature of institutional kinds and still have them. Here, obviously, it is important to be clear about what it means to be right or wrong about the nature of a kind, institutional or natural. Just to be able to identify a representative of a kind or to be able to apply a concept, one must have some knowledge and competence but that "some knowledge" may fall pretty far from knowing the nature or having a theory of the object.

And indeed, there is no debate about this, as Guala and Searle are both ready to accept this. Guala (2014, 62) quotes Searle saying the members of the community need not be consciously aware of the form of collective intentionality by which they are imposing the function on objects, and which according to Searle is a necessary condition of institutional kinds. And again, Searle says in the same quote that the members of a community can have false beliefs about the nature of institutional objects, their collective acceptance or imposition of a function may be based on a false theory.

So what is the problem or disagreement here? According to Guala, no collective acceptance or knowledge of the nature of the institutional object is required, and Searle agrees. However, as to collective acceptance of something as something, Searle explicitly requires it as a necessary condition. Here Guala seems to claim that the only option left for Searle is to employ a weaker notion of collective acceptance, and this appears to be on the right track. However, Guala's idea of the alternative weaker notion of collective acceptance is so weak that it cannot do any serious work anymore. Is this a problem – a false dichotomy? At least it seems that Guala's argument presupposes that the alternatives as to the notion of collective acceptance are (a) a very strong collective acceptance of "right theory about the nature of the institutional kind" or (b) a very weak notion of collective acceptance which boils down to an implicit agreement of the behavior with a rule, the conditions of which actually can satisfied by any pattern of behavior or any practice.

Here we think that there is an option between (a) and (b). This option is tracked by the method of Searlean transcendental reasoning or inference to the necessary explanation, which aims to answer the question of what must be true about the members of a community which is able to use, say, certain kinds of pieces paper as a means of exchange, that is, being able to use certain kind of objects as having a function they do not have barely on the basis of their physical properties. Here we think that Searle's train of thought goes along the following lines. First, the members of the community must use the objects as if they had the function in question. Second, they must do this intentionally, as mere behavioral regularity does not suffice for making sense of the practice. This requires that the members of the community must be able to represent the objects as having that function. This cannot be done only via the representation of the object as the physical object which it also is. Hence, to be able to use certain objects as having a certain function the members of the community must be able to represent the object as having the function in question. On this basis, the following kind of enabling condition of having institutional facts or an institution in a community can be identified: the members of a community must collectively accept that certain kinds of objects have a certain kind of function for them. This collective acceptance is attributed to the members of the community by the theorist, it is expressed in the behavior of the community. However, it is not ascribed to the members of the community as explicit thought contents in their heads, as they are not necessarily aware of this, as pointed out above.

Now the question is whether it is a necessary condition of the practice the members of the community maintain but yet not known to them, how is it possible that such a condition is doing any causal work in the existence or maintenance of the practice? Here we understand that Searle's answer, considered unsatisfactory by many, is that the collective acceptance is in the facilitating background of the intentional states of the members of the community. According to Searle's argument, we need to postulate such beliefs into the background to be able to account for the capacity to see the social world as we do, and make the actions of treating something as money for instance possible. These are no way conscious beliefs of agents, they do not understand they are doing this when they are treating Xs as Ys. Indeed their conscious beliefs and understanding of the social world can be crudely mistaken. What really makes the institutional entities have the function they do, is that they are treated in a certain way. This behavior is causally dependent on the representations of the agents, and the representations playing the causal role presuppose the facilitating constitutive beliefs in the background.

As the second step of his argument against necessity of collective acceptance Guala (2014, 62) continues: "I will argue that the conditions C that are

usually taken to be necessary for kindhood are actually redundant and can be dispensed with. For this reason, what people believe or accept regarding K is not constitutive of institutional reality." As already pointed out Guala uses C in a different sense from Searle's usage. Here, in Guala, C is used roughly to express the conditions X must satisfy to count as Y, which is not in the same way explicit in Searle's account (see formula 2). In Searle's account, these conditions are thought to be conventional and left open. Even though they can be pretty much anything, there must be some conditions that X must satisfy in order to count as Y in the community in question (as in formula 1), in part for epistemic reasons. In Guala's account, the condition C plays the role of a coordination device, they facilitate the beliefs required for the behavior in equilibrium (the solution of a coordination problem).

Guala asks the following question: "Why do we accept worthless paper bills in exchange for valuable goods and service?" This is a tricky question, because it seems to be in need of a corrective move due to description sensitivity. If we set the question this way, the answer to Guala's question is, actually, that we do not accept worthless pieces of paper in exchange for something valuable, rather we accept pieces of paper on which the function of medium of exchange has been imposed (and this makes them valuable). It is true that this function could have been imposed on many other kinds of entities instead of pieces of paper.

The collective acceptance, "for us," move in Searle's account guarantees that we all accept the same kind of things as having the function which in turn facilitates the coordination, and makes it rational for the members of the community to offer such pieces of paper as medium of exchange and believe that others will accept the pieces of paper as such. Indeed, for Searle, collective acceptance involves more than expectations, namely, commitments to treat Xs as Ys (Searle 2010, 82-86; see also Tummolini and Castelfranchi 2006, 314).

The role of the Central Bank relates to the well-functioning institution of money in a rather well-developed society having such high-level institutions like the Central Bank. It is not, however, central in the discussion of the basic ontology of institutions or institutional facts. Here it is important to distinguish between questions concerning features and properties of well-functioning institutions of a modern society and questions concerning the very basic features making the existence of institutions and institutional facts possible in the first place. It seems that the first kind of questions are issues to be studied by relevant empirical sciences, and the latter possibly suitable questions for a philosophical study. So the questions concerning the proper functioning of the institution and the why-questions concerning the behavior and reasons for it are empirical and scientific, the "how possibly" question is a philosophical one. The relation between these two kinds of question sets seems to us to be

that the why-questions really make sense when we have answered the "how possibly" question. It seems to us that Searle's focus is on the very rudimentary possibility of institutional facts, and his account does not aim to say anything about the conditions of properly functioning institutions. It seems that what Guala says in his argument to the effect that C conditions (in Guala's sense) are redundant and can be dispensed with is compatible with Searle's account, and does not make his argument against the necessity of collective acceptance any stronger.

Conclusion

In this article, we have identified several alternative ways to block Guala's argument according to which collective acceptance theories of social institutions lead to antirealism and infallibilism. One premiss in Guala's argument is that real kinds are not compatible with ontological dependence. We have three options to address this premiss: The first is that we bite the bullet, we accept that indeed the Searlean collective acceptance account presupposes the ontological dependence as characterized by Guala and yet claim that it is indeed possible to have both real kinds and such ontological dependence. Here the claim would be that the correct interpretation of ontological dependence in Searle's account does not have the unintuitive consequence that we create causal relations in the world by simply believing they exist. Defending this may require quite a bit of argumentative agility, however. The second way to face the premiss is indirect. It says that as to the ontological dependence, interactive kinds and institutional kinds stand and fall together. Guala has no qualms about interactive kinds being real kinds, so we can try to argue that he should either also accept institutional kinds as real kinds or exclude interactive kinds from the class of real kinds. The third option, perhaps the most promising one, is to argue that Guala reads Searle inaccurately and under the right rendering Searle's account does not presuppose ontological dependence either in the case of interactive kinds or institutional kinds. Here the steps of reasoning would be the following ones: (a) the charitable reading of Searle's account supported by textual evidence is that Searle is strongly committed to institutional kinds (institutions, institutional facts) being real. One can say this despite the unfortunate vagueness in some of Searle's formulations. Searle's project is clearly to accommodate institutional facts into a "naturalistic ontology": "We live in exactly one world, not two or three or seventeen" (Searle 1995, xi). He is rather obviously telling a story that is supposed to make room for such things as money, ministers, marriages in the scientific world view broadly taken. (b) We argue that there is no need for the problematic ontological dependence in Searle's story, the action is primary

and the role of representation is causal, however, representation is conceptually prior because it is required for the kind of behavior that causally brings about the institutional reality. The important element and the point of possible disagreement is the necessity of collective acceptance. Our claim is that collective acceptance is needed, maybe in a somewhat watered down sense by Guala's standards but yet needed, and indeed entailed by Guala's own account as well. (c) The outcome will be that Guala and Searle agree upon most of the elements of institutional reality but Guala needs collective acceptance in his account to get the institutional edifice up and running.

Declaration of Conflicting Interests

Funding

References

Aydinonat, N. Emrah, and Petri Ylikoski. 2018. "Three Conceptions of a Theory of Institutions." *Philosophy of the Social Sciences* 48:550-68.

Boyd, Richard. 1991. "Realism, Anti-foundationalism and the Enthusiasm for Natural Kinds." *Philosophical Studies* 61:127-48.

Glock, Hans-Johann. 2012. "Strawson's Descriptive Metaphysics." In *Categories of Being: Essays on Metaphysics and Logic*, edited by Leila Haaparanta and Heikki Koskinen, 391-419. Oxford: Oxford University Press.

Guala, Francesco. 2014. "On the Nature of Social Kinds." In *Perspectives on Social Ontology and Social Cognition*. Studies in the Philosophy of Sociality, edited by Mattia Gallotti and John Michael, vol. 4, 57-68. Dordrecht, The Netherlands: Springer.

Guala, Francesco. 2016. *Understanding Institutions: The Science and Philosophy of Living Together*. Princeton: Princeton University Press.

Guala, Francesco, and Frank Hindriks. 2015. "A Unified Social Ontology." *The Philosophical Quarterly* 65 (259): 177-201.

Hacking, Ian. 1999. *The Social Construction of What?* Cambridge: Harvard University Press.

Mäki, Uskali. 2005. "Reglobalizing Realism by Going Local, or (How) Should Our Formulations of Scientific Realism Be Informed About the Sciences?" *Erkenntnis* 63:231-51.

Searle, John. 1995. *The Construction of Social Reality*. New York: The Free Press.

Searle, John. 2010. *Making the Social World*. New York: Oxford University Press.

Strawson, Peter F. 1959. *Individuals*. London: Methuen.

Thomasson, Amie. 2003. "Realism and Human Kinds." *Philosophy and Phenomenological Research* 67 (3): 580-609.

Tummolini, Luca, and Cristiano Castelfranchi. 2006. "The Cognitive and Behavioral Mediation of Institutions: Towards an Account of Institutional Actions." *Cognitive Systems Research* 7:307-23.

Tuomela, Raimo. 2007. *The Philosophy of Sociality*. Oxford: Oxford University Press.

Tuomela, Raimo. 2013. *Social Ontology: Collective Intentionality and Group Agents*. Oxford: Oxford University Press.

Author Biographies

Pekka Mäkelä is a philosopher interested in the philosophy of the social sciences, and philosophy of collective and social action, in particular their normative dimensions. Pekka Mäkelä is the Head of Discipline of Practical Philosophy at the Faculty of Social Sciences, and the coordinator of the Centre for Philosophy of Social Science (TINT) located in the discipline of Practical Philosophy at the Faculty of Social Sciences of the University of Helsinki.

Raul Hakli is a university researcher in Practical Philosophy in University of Helsinki, Finland. His research interests include collective intentionality, social ontology, and social epistemology.

S.M. Amadae is a lecturer in World Politics in Political Science at the University of Helsinki; associate professor of International Political Economy at Swansea University, Wales; and research affiliate with Science, Technology, and Society at the Massachusetts Institute of Technology. Amadae is the author of *Rationalizing Capitalist Democracy* (University of Chicago Press, 2003) and *Prisoners of Reason: Game Theory and Neoliberal Political Economy* (Cambridge University Press, 2016), and most recently "Game Theory, Cheap Talk and Post-Truth Politics: David Lewis vs. John Searle on Reasons for Truth-Telling," *Journal for the Theory of Social Behavior* (2018).