John Searle’s ontology of money, and its critics


John Searle has proposed one of the most influential contemporary accounts of social ontology. According to Searle, institutional facts are created by the collective assignment of a specific kind of function — status-function — to pre-existing objects. Thus, a piece of paper counts as money in a certain context because people collectively recognize it as money, and impose a status upon it, which in turn enables that piece of paper to deliver certain functions (means of payment, etc.). The first part of this essay presents Searle’s theory of social ontology and summarizes his views on money. The second part centers on criticisms of his theory and compares his theory with its main competitors, including the theories of Tuomela, Guala and Hindriks, and Smit, Buekens and Duplessis. The conclusion attempts to see what may remain of Searle’s theory after careful examination.

Keywords for index: John Searle, constitutive rules, regulative rules, institutional fact, brute fact, collective intentionality, incentives, deontic powers, realism, naturalism, status-functions

By
Louis Larue
Researcher in Practical Philosophy
University of Gothenburg, Sweden
louis.larue.2@gu.se.

Draft manuscript
Submitted August 1, 2022

Acknowledgments: I would like to thank Ragnar Francén, Anna Sofia Maurin, Scott Scheall, Joe Tinguely Maximilian Zachrau and Sebastian Östlund for their insightful comments and suggestions on this paper.
Introduction

John Searle has proposed one of the most influential contemporary accounts of social ontology, whose main aim is to determine rigorously the nature of institutions. Searle’s social ontology attempts to explain how we routinely create institutional facts, such as money, political boarders, debts and laws. According to Searle, institutional facts are created by the collective assignment of a specific kind of function — status-function — to pre-existing objects, persons or state of affairs. Thus, a piece of paper is money in a certain context because people collectively recognize it as money, and impose a status upon it, which in turn enables that piece of paper to perform certain functions (means of payment, etc.).

All these terms and many more require careful definitions and examination. The first part of this essay presents his theory of social ontology and summarizes his views on money. For money has been one of Searle’s favorite examples, and he has recently further developed his thoughts on that matter. Moreover, Searle’s account of the ontology of money has been criticized by many and has led to fruitful debates between economists and philosophers. The second part of this chapter reviews the most important criticisms and discusses their relevance. The conclusion attempts to delineate what may be the relevance of Searle’s theory for the social sciences.

Searle’s theory

Some preliminary distinctions

A theory of institutions may have several purposes. One could be to explain the historic origin of specific institutions. For instance, Douglass North proposes a complex account of how a certain pattern of bargaining between political and economic actors explains the emergence of specific property rights in the Western world. Another purpose could be to explain the stability and persistence of these specific institutions. For instance, North shows how interests’ lock-ins sometimes prevent institutional change. A further, third possible goal is to explain why there are institutions at all. The answers of most authors resort to functionalism (of various sorts): institutions exist because they fulfil a function. North, again, explains their existence by their role, which is to reduce uncertainty and transaction costs, though he does not say clearly whether that function is assigned intentionally or not. The case of money is no exception: most economists generally hold that its existence is explained by the fact that it facilitates trade. This view is also popular among philosophers, such as Guala and Hindriks: according to them, institutions exist because they generate “cooperative benefits”. I shall later on discuss how functions fit in Searle’s account, as well as the precise meaning he gives to that concept. As we shall see, though, both its role and its definition differ sharply from Guala’s, Hindriks’s, and North’s accounts. A fourth, related question, concerns the efficiency of institutions, that is, how well they serve their function. This question is central to North’s account, though Guala, Hindriks, and most philosophers with an interest in social ontology leave it aside.

4 North, Structure and Change in Economic History; North, Institutions, Institutional Change and Economic Performance.
5 Tobin, ‘Money’.
The primary aim of Searle’s theory lies elsewhere.7 His main goal is to inquire into the nature of institutional facts.8 Searle takes for granted that institutional facts exist within the physical world. As he claims repeatedly, their existence is so obvious that they become “invisible”; we fail to see their specific features, and to many of us they seem as natural as stones, water and air.9 His project is to uncover the nature of these institutional facts. He also wishes to solve an apparent puzzle: we say that money, debts and boundaries exist. Yet they seem to exist merely because we believe or say that they exist. How can we make sense of these claims? Searle starts out his inquiry by making two important distinctions.

First, some features of the world are observer-independent while others are observer-dependent. If there were no conscious human beings with intentional states on this planet, mountains, rocks, and seas would continue to exist (they are observer-independent features of the world), while legal constitutions and currencies would not (they are observer-dependent features). Yet, Searle argues that all these things are, in some sense, objective. Hence he makes a second distinction between the ontological and the epistemological sense of the subjective-objective divide. The epistemological sense refers to judgments about statements. The statement “copper is cheaper than gold” is epistemologically objective: it can be true or false depending on some objective facts (the price of each product). The statement “gold is more beautiful than copper” is epistemologically subjective: it is a matter of personal judgement. The ontological sense refers to modes of existence. Something is ontologically subjective if it requires experience by a sentient being for its existence. Something is ontologically objective if it does not. Pain and pleasures are prime examples of ontologically subjective things, mountains and rocks are examples of ontologically objective things.10

Thus Searle is primarily interested in explaining the nature of (a specific kind of) observer-dependent features of the world, which we create through our subjective attitudes, and about which we can make epistemologically objective statements.

Finally, it is important to stress Searle’s realism and naturalism. Searle is a realist (or, to be more precise, an external realist) that believes that there is a physical reality out there, independent from us and of our representations of it.11 He is also a naturalist that considers that the capacity of human beings to create institutional facts is an extension of more basic biological phenomena, such as our capacity to cooperate.12 Therefore, for Searle, the institutional realm is not a distinct metaphysical realm. Contrary to some interpretations, Searle does not argue that there is a “social world” entirely distinct from the “physical world”, or that there is a specific ontology of the social world distinct from the ontology of the physical world.13 His purpose is rather to derive the institutional world from the physical world.14 Let me take an example: a bill has certain physical properties: it is made of paper, for instance. But it has also certain non-physical properties that are assigned to it by human beings: it may be used to buy things, or to light a fire, among other things.
However, there is not a “bill the physical object” and a “bill the social object”. The bill remains one and the same object. In other words, as we shall see in the next section, even if one can say that it is a brute fact that a bill is made of paper, and an institutional fact that it may be used to buy things, the bill always remains one and the same object.\textsuperscript{15}

Searle’s inquiry focuses on institutional facts, which, as we shall see very soon, are a sub-set of social facts. His main questions, then, are: what are institutional facts? How do they relate to brute facts? How can we explain the institutional world as an extension of the physical world? Or, in the context of this chapter, how does a piece of paper with certain physical features become money?

**Institutions and institutional facts**

Let me briefly state Searle’s theory, before going into the details. Searle defines an institutional fact as constituted by the collective assignment of a status-function upon a pre-existing object, person or state of affairs in a certain context. The creation of institutional facts by assigning status-functions also comes with the creation of deontic powers. Moreover, he claims that institutional facts can only exist within an institutional framework. That distinguishes them from brute facts\textsuperscript{16}, which do not require a broader institutional framework. The institutional framework sets out the rules that enable the creation of institutional facts. All these terms may seem esoteric. Let me explain each of them in turn.

Searle assigns great importance to collective intentionality in his account of institutions.\textsuperscript{17} He says repeatedly that “institutions are collectively accepted systems of rules” and speaks of the “collective assignment of functions”.\textsuperscript{18} We generally think of intentions as personal: “I intend to do this” “she intends to do that”. How can we make sense of “we-intentions”? Let me start with an example.\textsuperscript{19} When we say that two people are walking together in the same direction, we might mean, first, that we observe that they both happen to walk side by side, each intending to go in that direction, but with no collective intention of doing so. They are “alone together”, as Butchard and D’Amico nicely phrase it. This shows that joint action is not sufficient proof of collective intention. Rather, when we say that two persons are walking together as part of a collective intention, we mean that they each take part in the joint action and share the joint intention. “My” doing this action is part of “our” doing this action.

According to Searle, collective intentionality is a common feature of both social facts and institutional facts. As an example of the former, take a mob of hooligans collectively intending to storm the president’s palace. Or a herd of wolves chasing a sheep. Though similar in structure, institutional facts involve more than collective intentionality. What sets them apart from social facts, according to Searle, is the assignment of a specific kind of function, which he calls status-function.

\textsuperscript{15} e.g. Smith and Searle, ‘The Construction of Social Reality’, 300.

\textsuperscript{16} That distinction goes back to Anscombe (‘On Brute Facts’) and Searle’s earlier work (Searle, Speech Acts. Note that brute facts require the institution of language to name them, but their existence does not require it. See Searle, The Construction of Social Reality, 27.

\textsuperscript{17} Unfortunately, Searle’s definition of collective intentionality is somewhat imprecise. In particular, he remains obscure as to how, in practice, something is collectively recognized as an institution. Fortunately, however, Searle is not the only philosopher to argue for the existence of collective intentions. Interested readers can benefit from the work of Raimo Tuomela. For instance, Tuomela, The Philosophy of Social Practices; Tuomela, The Philosophy of Sociality.


\textsuperscript{19} Drawn from Butchard and D’Amico, ‘Alone Together’.
Physical objects do not have an intrinsic function in virtue of their physical properties. Rather, Searle claims, they always acquire a function because humans (or other conscious animals with intentional states) assign it to them. The fact that the heart pumps blood into the body does not entail that it is its function. Humans say that it is because they value life over death, and good health over illness. They assign life-maintaining functions over it, which it does not possess “naturally”. In short, functions are always observer-dependent. Searle claims that some of these functions are “non-agentive”, such as the heart’s function to pump blood, because these functions are simply part of a theoretical account of the object in question. Others are “agentive”, because they are assigned to objects by human agents to serve a practical purpose of theirs. For instance, the agentive function of a knife may be to cut meat, or to kill (the original function need not be the only one). Finally, among agentive functions, there is a special class of functions, which Searle calls “symbolic”: the function of representing something else. Thus, for instance, the function of raising my hand may be to say that I have a question. Similarly, a line on the ground (or a river) may stand for a border between two countries. Finally, these functions may be imposed on objects without anyone consciously thinking about it. Similarly, the historic moment when the original function was imposed may well be forgotten. We may be so used to the fact that the Channel is the border between Great Britain and Europe that we forget that it is a symbolic and agentive function that was once imposed on it.

When you combine the assignment of functions with collective intentionality, you get the possibility for the collective assignment of functions. So a group of people can collectively recognize a piece of wood intertwined in a piece of metal as a hammer and grant it the function of breaking ice, among other things. Yet a hammer is not what we usually call an institution. The missing piece is the assignment of status-function.

Status-functions are a special kind of agentive symbolic functions. As the name suggests, objects to which a status-function is assigned acquire a status. And this status allows them to perform a function. For instance, the object (or speech act) “habemus papam” is given a status that allows it to turn people into Popes, if it is uttered by certain people in a certain context. Importantly, the object upon which a status-function is imposed cannot perform its function solely in virtue of its physical properties, but in virtue of the fact that there is a collective assignment of status to that object. So “habemus papam” serves its functions not simply in virtue of the sounds it produces. In addition, the status may well survive the demise of the object: a pope remains a pope after “habemus papam” has been uttered and it need not be reasserted repeatedly.

In short, Searle uses the formula “X counts as Y” to describe the logic of status imposition, where X is the object and Y the status. However, certain objects acquire a status only in certain contexts. “Habemus Papam” turns someone into a Pope only on certain occasions in a specific place. Thus, Searle uses the more specific formula of “X counts as Y in C”, where C refers to the context. One of Searle’s favorite example of status-imposition is money: a piece of paper with certain physical traits (X) counts as a dollar bill (Y) in the USA (C). Money, for Searle, is created by the collective assignment of a status-function upon a certain object in certain contexts.

---

21 Searle, 20.
The collective assignment of status-functions to objects always comes with the collective assignment of deontic powers to certain persons or groups. These may be rights, obligations, duties, permissions, requirements and prohibitions. When we collectively count an X (a pre-existing object) as a Y (a status) in some context C, we recognize a certain distribution of deontic powers over agents of various kinds. Deontic powers are essential, for Searle, because the function that comes with the imposed status cannot be performed without them. Money, for instance, can play its role as a means of payment only if it comes with adequate powers, such as being able to pay with it. When we count a piece of paper as a dollar bill, this assignment of status gives its holder the power to buy stuff, to pay her debts and taxes, to make other people do certain things for her against payment, etc. Similarly, the recognition of a river as a border means that some people are allowed to cross it while others may not. Certainly, the river does not have these powers because of certain physical traits. Though certain of its physical traits do constrain our freedom of movement, counting it as a border comes with additional powers, which are explained by its assigned status.

We have finally arrived at Searle’s definition of institutional facts, which are created by the collective assignment of a status-function upon a pre-existing object in a certain context. Moreover, the creation of institutional facts also creates deontic powers of various sorts.

The picture of institutional facts that comes out of this short description may give the impression that these are free-standing facts, independent of each other. That is of course not the case. First, institutional facts relate to each other in various ways. That a piece of paper with certain physical traits is a dollar bill is an institutional fact. But its existence depends on the existence of other institutional facts, such as the United States of America, the Federal Reserve, the Treasury, etc. In other words, institutional facts form a network, and make sense only within that network.

Second, these institutional facts, Searle claims, require specific institutions for their existence. They are embedded in an institutional structure, that allow for the creation of institutional facts. For instance, the institutional fact “contract” requires the institution of “property”. These statements may seem circular. Let’s try to make them as clear as possible. How do institutional facts relate to the institutional structure (or to “institutions”, in short)? How does “contract” relate to “property”?

For Searle, institutions are systems of constitutive rules. He contrasts constitutive rules with regulative rules, a distinction that originates from Rawls’ early work and that is central to Searle’s account. Constitutive rules create the possibility for the mere existence of certain activities. Searle’s favorite example is the rules of chess. Chess would not exist without these rules. Regulative rules, on the contrary, are of the form “drive on the left side of the road”, or “eat with your right hand”, or “write your exam with a pen”. These rules regulate the activity of driving, eating and examining, respectively, but they do not create these activities and are not constitutive of them. You can drive, eat or examine without following these rules. In other words, regulative rules merely

23 Of course, something cannot count as money solely because it is recognized as money. It also should function as money. See Khalidi, ‘Three Kinds of Social Kinds’. As Guala (2010, 260) writes: “What counts as money does not depend merely on the collective acceptance of some things as money, but on the causal properties of whatever entities perform money-like functions.” This is why the imposition of status-functions, and the correlated deontic powers, is so important for Searle’s account. Something counts as money if it acquires some functions through the collective assignment of a status.

define conventions, while constitutive rules are essential to institutions. The difference between the two is that constitutive rules make possible institutional reality (e.g. rules of chess), while regulative rules define conventions that govern that institutional reality (e.g. color of the pieces).

We have already encountered constitutive rules, without knowing it. For Searle argues that constitutive rules are of the general form “X count as Y in C”, where X is some object, person or state of affair, Y is a status imposed on that fact, C is the context, and “count as” indicates the assignment of status-function through collective intentionality.

Searle claims that what underlies all human institutions is the capacity to impose status-functions upon objects according to constitutive rules with the former logical structure. The imposition of status-function thus requires the existence of a background structure of constitutive rules. For instance, the fact that the Green party has lost the election requires electoral rules (or the institution of election). The fact that I sign a contract with a company for the delivery of some goods requires rules defining property rights (or the institution of property). Without these rules, the facts of electoral defeat and of signing a contract make no sense. To repeat, the creation of institutional facts is a creation of something new through the assignment of status to an object, person, or something else. This is made possible by the existence of constitutive rules. Hence, institutions, for Searle, are systems of constitutive rules that enable the creation of institutional facts.

Finally, all institutions and institutional facts require language. As can be expected from a philosopher of language like Searle, he dedicates a lot of attention to this question. I will not delve too much into that part of his theory, and only focus on the most important bits. Institutional facts are a matter of status-functions, and status-functions need to be represented in order to exist (as we have seen before, status-functions are a special kind of symbolic functions). For Searle, a piece of paper acquires the status “dollar bill” only in so far as it can be represented as a dollar bill, and this requires language (or symbolism in a larger sense). A dog who looks at a dollar bill only sees the paper, the ink, etc. He does not see the bill as a piece of money. On the contrary, when I see a dollar bill, I do not simply see a piece of paper with some markings on it. I also see the piece of paper as a dollar bill, with the status and the deontic powers that come with it. This is because I am able to see the symbolic functions that it carries, thanks to the existence of language as a medium of representation (which the dog lacks). In short, institutional facts exist only because people recognize and accept that they exist, and this requires representation and thus language.

**Criticisms**

Searle’s account of institutions has attracted a lot of criticism – too much for a single chapter. Here I mostly concentrate on the debate about his views on money. First, I consider whether explaining the nature of money requires the existence of physical objects upon which to impose functions. Second, I consider how theories on the nature of money relate to how effectively money fulfils its functions, and whether these normative issues actually matter, or could be disregarded. Third, I discuss whether collective intentionality is really needed to account for institutions in general, and money in particular. These three specific objections tackle Searle’s view on money. Yet they also will enable me to draw wider conclusions regarding the purpose of social ontology, and the benefits of different approaches. I then move on to studying two broader criticisms of Searle’s account.

---

These objections do not apply to money in particular, but, if successful, they would threaten Searle’s entire edifice and thus weaken the relevance of Searle’s project for discussing the nature of money. Moreover, a review of the debate on Searle’s work would not be complete without a discussion of these issues, which have proved central in the debate on his theory. The first general objection concerns the regulative/constitutive rule distinction. The second relates to the reasons for which people stick to specific institutions.

Object

Searle claims that his account is particularly well suited for describing money. Yet, an obvious problem was soon pointed out by Barry Smith in an exchange with Searle. Searle’s account assumes that institutional facts always rely on (or are imposed upon) physical objects or persons or state of affairs. That is fine for bills and coins, but what about bank accounts and Central bank’s reserves? Should we say that lines on bank records or electronic “blips” on electronic records are money? There is an apparent lack of a “X” term here, and the “Y” term seems to be standing on his own.

Searle actually came to agree with Smith. He now shares the view that we are moving towards a world without physical currency, and so very often there is no physical object on which a monetary status is imposed. We have records of money, but no hard money. His answer is that what really matters, in the last instance, is not the object on which status is imposed, but the deontic powers that are bestowed on people through status-imposition. So, for instance, “my having a thousand dollars is not a matter of my having a wad of bills in my hand but my having certain deontic powers.”

Or, to put it differently, it is possible, for Searle, to conceive of deontic powers being bestowed upon people without any physical intermediate. Money, then, is simply power, i.e. the “deontic power to buy, pay and close debts.” Many other philosophers have defended a similar claim. G.A. Cohen, for instance, claims that money is not primarily an object, but a power, that is, purchasing power. It can be solely present in the mind, without any need for material equivalents. In a similar vein, Smit, Buekens, and du Plessis argue that bills, coins and bank records are all merely records. What really matters is the purchasing power that they represent.

However, I would like to argue that Searle’s original account may be fine as it is, and that he should not accept Cohen and Smit et al.’s objection too quickly.

First, contrary to what Barry Smith argues, bank records, whether electronic or material, have a material structure. They are stored on very solid computers and servers. These are not free-standing stuff. Second, bank records are money, similarly to coins and bills. Some claim that they are not because, if a server or record is destroyed, money does not disappear. Similarly, if a bill is damaged, it is usually replaced by another one by the Central Bank. Therefore, money is not a material object, but a form of power that sometimes but not always is attached to a material object.

In reply, I shall note, first, that “damaged” is not the same thing as “destroyed”. If a server is damaged, you may well be able to retrieve what is stored on it. That’s hopeless if it is destroyed.

---

29 Smith and Searle, 307.
31 Searle, 16.
33 Cohen, On the Currency of Egalitarian Justice and Other Essays in Political Philosophy, 185n39.
35 Smit, Buekens, and du Plessis, 18.
Similarly, money does disappear if all records of it are destroyed. There are plenty of examples of this. Many people lost their Bitcoins when the MtGox online exchange platforms crashed, and these Bitcoins disappeared. Similarly, people may irremediably lose their emails, their PhD thesis, their movies, or whatever, if the server on which it has been stored crashes. Of course, banks usually keep copies of their (electronic) records, which reduces the risk of loss. But if all banks lost all their records, money would simply disappear, in a very similar way that if you burn a 100-dollar bill, it disappears.

One may reply that, contrary to bills, making a hundred copies of your bank records does not multipy the value of your deposit by one hundred. That’s the point of Smit et al.’s argument: records are mere records, they are not money. To make that point clearer, let me take a non-monetary example. We usually sign 2 or 3 copies of our employment contract. Yet there is only one employment, not 2 or 3. Thus, taken literally, Barry Smith’s critique of Searle’s conception of money may well extend to all institutional facts. In some sense or another, all institutional facts have the property of being a record of something else. Money is not peculiar at all in that respect. Is the entire edifice threatened? I do not think so. The apparent threat is in fact the product of a confusion.

The entire point of Searle’s approach is to highlight that institutional facts are facts that point at something else. In Searle’s jargon, they are facts that count as something else. I think that Barry Smith, G.A. Cohen and others are all running into an important confusion here. They confuse the fact for its function and its deontic powers. For instance, the fact that something is a dollar bill should not be confused with its function as a means of exchange and with its purchasing power. Or the fact that this sheet of paper is a contract is not equivalent to its function of signaling that I agree to work for you. “dollar bill” and “purchasing power” are inherently related of course. Yet each operate at a different level. The ground level is made of physical objects (piece of paper, electronic blips on servers). Then there are institutional facts (dollar bill, bank records). Then, at the highest level, there are functions and deontic powers such as “purchasing power”. As G.A. Cohen argues, there is no need for a written contract for me to know that I owe you a certain amount of money. We may well one day engineer human brains so that direct telepathic communication and secure debt records within brains will render all these material things obsolete. Yet, even in that fantasy world, there will be a difference between the concepts of ownership, my ownership of a certain amount of money and the fact that I store this specific ownership-relation on brain cells.

Barry Smith also pointed out a further problem related to the object on which a status is imposed. Not only could there be “objectless” institutional facts (Y without X), but people can also wrongly count an X as a Y. Smith takes the example of counterfeited money. Suppose a counterfeiter is so successful that people cannot see the difference between fake dollar bills and true ones. A function is then wrongly, or mistakenly, imposed on an object. Counterfeiting and other forms of fraudulent behavior are an important concern for law enforcement authorities. Searle’s account, however, is not normative: his purpose is not to say what institutions should do, how they should be designed, or what are the “true” or “good” or “adequate” institutions. Rather, his goal is to describe their nature and explain how they are created through collective intentionality. Fraudulent institutions, or counterfeited money, are institutions, and they can be accounted for by pointing to their status-

---

functions and deontic powers. There are of course constant disputes on the design of just institutions, but that is the business of political philosophy, not social ontology. I shall come back to similar misunderstandings in the next subsection.

Functions and efficiency

A related, though different, issue was pointed out by several authors, including Barry Smith, Francesco Guala and Frank Hindriks. Very often, money fails to fulfil its functions, or does it in an inefficient, or incomplete, or somehow inadequate way. Again, that is a real problem for normative theory – but is it a worry for social ontology? Smith, Guala and Hindriks do think so: a proper theory of institutions ought to explain not only the nature of institutions but also their success or efficiency, which Searle’s theory fails to do. Searle disagrees with the latter claim. He makes clear that his theory is not normative and argues that the normative assessment of the efficiency or fairness of institutions falls entirely outside its scope. Even if his theory entails that the creation of institutional facts comes with the creation of deontic powers, he refrains from discussing whether these powers are fairly distributed, or efficiently put to use.

Searle’s stance is controversial. In economics at least, it is often argued that what money is relates to what money does and how well it does it. As a correlate, it is argued that those currencies that fail to live up to certain standards should not be conceived as money. Thus, for instance, Yermack argues that Bitcoin is not a real currency because it does not fulfil the traditional functions of money sufficiently well. Similarly, Smit, Buekens and du Plessis claim that money is whatever is efficiently used as a means-of-exchange, that is, as a device to reduce transaction costs. Since Bitcoin is not largely and efficiently used as such, they argue that it is merely a financial asset.

I think that these claims are mistaken. The main issue is that it is very difficult, and perhaps impossible, to settle on a consensual standard of “good” money. Money’s functions are easy to define, but difficult to assess in practice. In fact, no currency could be said to fulfil all functions well for everybody all of the time. Does that mean that no currency can properly be called money? Are we all wrong to call dollars “money” after the 2007 financial crisis? Certainly not. But then, how well should a currency function, and for how many people, in order to count as money? Smit et al. argue that Bitcoin should not count as money because “at present, the vast majority of bitcoins are traded as a speculative investment, not as a means to lower transaction costs.” I wonder why this would disqualify Bitcoin as money. Large quantities of dollars, euros and pounds are also traded as assets, and speculation on currencies is a profitable business. Moreover, Bitcoins do serve as means-of-payments, though maybe not the majority of them. Finally, some currencies, such as local currencies, are used only in a very tiny geographical area. Yet, most of their users would say they are money. In short, we should not confuse what money is with how successful it is (whatever this means).

38 Smith and Searle, 301.
44 Smit, Buekens, and du Plessis, 333.
More generally, I think that squeezing into the definition of institutions an account of their efficiency just makes things more complicated: most institutions are ineffective in some way. Institutional ineffectiveness does not prevent us from counting them as institutions and does not create any serious problems for Searle’s account, contrary to what some may think. \(^46\) I shall come back to this point when discussing Guala and Hindriks’s own theory.

**Collective intentionality**

A far more powerful criticism of Searle’s theory relates to his use of collective intentionality\(^47\). Several authors argue that collective intentions are an unnecessary metaphysical commitment. Whether or not there are collective intentions out there, they argue that one does not need them in order to account for institutional facts. This is a strong objection, because Searle’s theory relies on collective intentionality as a bridge between physical objects and the institutional world. Thanks to collective intentionality, status-functions can be assigned to objects. Hence, if we can get rid of collective intentionality, we can also get rid of the entire apparatus of status-functions and constitutive rules. Consequently, because they either deny or discard collective intentionality, both Smit, Buekens and du Plessis, and Guala argue that institutional facts can actually be reduced to brute facts. \(^48\)

I shall not review the literature on collective intentionality here, nor settle the case for whether or not it exists, or makes sense at all. \(^49\) Rather, I want to assess whether it is possible to give a satisfactory account of institutions such as money (which is the prime example of interest in this chapter) without collective intentionality. I will briefly analyze two such attempts, which are each part of a larger literature whose main effort is to build game-theoretic accounts of institutions. That literature started out with Lewis’s study of conventions \(^50\), and has had a fruitful life of its own in economics. \(^51\)

A first interesting account is the one of Smit and his co-authors, who argue that facts such as money, borders, or traffic lights can be accounted for in terms of actions and incentives only. \(^52\) So, for instance, a traffic light is a traffic light because agents are incentivized to act in a certain way (“stop at red”, “go forward if green”) when encountering it. Instead of Searle’s “X counts as Y in C” formula, they prefer their own “Subject S is incentivized to act in manner Z towards object X”.

Guala and Hindriks have developed a similar game-theoretic account of institutions. They argue that institutions are solutions to coordination problems. More specifically, they claim that institutions are “correlated equilibria of coordination games with multiple equilibria”. \(^53\) A coordination game is a game in which several agents (say two) must decide what action to undertake (their “strategy”), and in which the outcome for each player depends on the other player’s action. An equilibrium of a coordination game is a profile of strategies (one for each player), where each player’s strategy is the best response to those of other players. Take the

---

\(^{46}\) Hindriks and Guala, ‘Institutions, Rules, and Equilibria’, 462.


\(^{49}\) See Schweikard and Schmid, ‘Collective Intentionality’.

\(^{50}\) Lewis, *Convention*.

\(^{51}\) For a review, see Binmore, ‘Game Theory and Institutions’.


\(^{53}\) Hindriks and Guala, ‘Institutions, Rules, and Equilibria’, 466; see also Guala and Hindriks, ‘A Unified Social Ontology’, 182–86.
example of two tribes having to decide where to hunt. There are two possible hunting grounds and both tribes cannot hunt at the same time in the same place. There are two equilibria in this game, for they can each hunt in either hunting ground. How will they agree on which land to hunt? (Suppose that they do not talk to each other). One way to solve the problem is to use correlation devices, that is, arbitrary “signs” or “pre-emption devices”, such as “whoever happens to be there first hunts first”. In technical terms, we say that a correlated equilibrium involves strategies that are conditional upon an event or signal sent by an external coordination device. For Hindriks and Guala, these coordination devices are “rules”. This is why they say that their account is a unification of the rules and equilibria approaches to institutions: these are simply “rules-in-equilibria”.54

These accounts, if successful, have several benefits. First, they do not seem to assume collective intentions. Nor do they suppose any kind of status-function imposition. The former view incentives as applying to individuals only, while the latter considers each player’s benefits in solving a coordination game. For both, institutions are simply a nice way to make the world better, by solving intricate coordination problems. These accounts are thus able to explain the existence of institutions by referring to their purpose, which, in a nutshell, is to reduce transaction costs and provide cooperative benefits to people55. Money seems to fit well within these accounts, for it does just that: reducing transaction costs and generating cooperative benefits.57

Let’s come back to our original question: do these accounts succeed in keeping collective intentionality aside? Their core claim is that institutions are just a matter of convention, which are the product of game-theoretic processes and which give us signals or incentives to act in a certain manner. In short, they claim that whatever object or rule happens to give proper signals for coordinating behavior, that is, whatever fulfils the function of coordinating behavior in a certain domain, is an institution. Thus, money is the institution that regulates the behavior of exchanging things. There is apparently no need for collective intentions for its creation.

In response, Searle argues that, contrary to collective intentionality, game-theoretic concepts are neither sufficient nor necessary to explain the nature of institutions.58 Game-theoretic equilibrium concepts are not sufficient to explain the nature of institutions because plenty of equilibria are not institutions at all. Thus, for instance, we may coordinate to row a boat at a certain pace. Yet there is no institution being created here. Similarly “raising your hand to ask a question” is not an institution. This is simply a convention, which might well be different without threatening the existence of the institution of the “seminar”. Moreover, the equilibrium concept is not necessary because plenty of institutions do not involve game-theoretic equilibria. Searle takes the example of property rights in the US, which are simply enforced and created by Federal institutions.59

Several elements of this debate need to be sorted out here.

---

54 Guala and Hindriks, ‘A Unified Social Ontology’.
56 Hindriks and Guala, ‘The Functions of Institutions’.
57 These accounts have problems of their own, which I will not review here. One important issue is related to their functionalism. My main worry is that they do not provide any empirical evidence for their claim that the existence of particular institutions is explained by their function. This is simply an unwarranted functionalist explanation. For a thorough consideration of these issues, see Elster, Jon. Explaining Social Behavior. (Cambridge: Cambridge University Press, 2007). Another worry concerns the alleged “benefits” of institutions. I wonder how anyone could argue that slavery, and other colonial institutions, can be said to benefit people in any meaningful sense.
58 Searle, ‘Status Functions and Institutional Facts’, 511.
59 Searle, 511.
First, to repeat, we should be very careful to distinguish how a specific institution has historically come into existence from the nature of this institution. Thus, some institutions are perhaps the result of government actions, while others are probably the product of correlated equilibria. This is a question for economic historians, which is independent from the questions raised by Searle. To put it differently, even if it were possible to give a theoretical account of every fact of human life based on game theory, this would be no proof that these facts are really the product of game-theoretic processes. Only the study of history can tell us what the origin of specific institutions is. Hindriks and Guala’s account, as well as Smit’s, are primarily theoretical, and remain vague on history. While the question of how certain institutions came into existence is central to their account, it has no incidence for Searle, whose main purpose is to give a general account of what institutions are, not of their origin.

Second, Searle contends that even if some institutional facts may well be the historical product of game-theoretic equilibria, that does not explain their nature. What distinguishes these “game-theoretic” institutional facts from other non-institutional “game theoretic” facts is that the former involve the collective imposition of status-function that creates deontic powers. There is no need to impose a status on rows in order to row together, but there is such a need for traffic lights, money, or borders. What Hindriks and Guala describe is a possible process through which a given status has been imposed on a given object. People may well have agreed to count “green” as meaning “go forward” and “red” as meaning “stop” through game-theoretic processes described by Guala and Hindriks (though frankly I doubt it). Yet what this means is that they have assigned a certain status to a certain object. And the status may remain even after the original imposition has been forgotten.

A third, related problem common to game-theoretic accounts of institutions is that they often presuppose what they are supposed to show. As Searle argues, they “presuppose the existence and nature of institutions and then try to give grounds for selecting an institution.” For instance, Smit et al. never explicitly say what gives rise to incentives. The object “traffic light” (their favorite example) cannot incentivize people just in virtue of its physical traits. Similarly, a dollar bill does not acquire purchasing power simply in virtue of its physical traits. As we have seen, functions do not arise naturally from objects. If that is true, what is it that creates the incentive to stop at traffic lights, or to pay and be paid in dollars? Clearly, something is missing from their account of institutions, namely, the institutional facts that explain why “green” means “go forward”, and “red” means “stop”. Without an explanation of the origin of incentives, their account implicitly presupposes their existence. Hence their account presupposes what it is supposed to explain.

A fourth fundamental issue with these game-theoretic accounts will bring us back towards our original purpose: are collective intentions necessary? Let us have a look at what Hindriks and Guala, Smit and his co-authors, actually describe. Whether it is the traffic light example or the hunting example, what we see through their theoretical eyes is a bunch of people making gestures in the same direction or in the same way. People moving or stopping their cars, chasing animals, exchanging pieces of papers, etc. In the words of Butchard and d’Amico, we see people moving “alone together”. In the world of Guala and Hindriks, and of Smit et al., each individual is doing

---

60 Elster, *Explaining Social Behavior*.
63 Searle, 513.
something with other people but with no collective intention to do so together (since there is no collective intentionality). This is hardly what we would call collective behavior. It would be rather puzzling to describe a collective hunt as made of separate hunters who merely happen to hunt on their own alongside fellow members out of their own interest and as a best response to the behavior of others. This may be part of the explanation for their behavior, but that cannot be all of it. What we do see in our everyday life is people collectively accepting certain colors in certain contexts as signs for stopping or going forward; and a happy lot of folks enthusiastically pursuing a hunt. Someone fails to grasp an important part of social and institutional reality if this is ignored.

We have thus reasons to think that collective intentions are necessary to account for institutional facts.

**Rules**

Yet the debate is not over. Hindriks raised the following additional and related criticism: constitutive rules can be derived from regulative rules, and have no separate conceptual existence.\(^{66}\) So the distinction between conventions and institutions that is so central to Searle’s theory is untenable. If true, this means that Searle’s claim that Guala and Hindriks, and other scholars informed by game theory, confuse the two is simply mistaken. All institutions are conventions.

Take, for instance, the case of property. According to Searle, a piece of land may become my property if people collectively recognize that this land (X) counts as my property (Y) in a specific legal context (C). Hindriks, and later also Guala, argues that this constitutive rule (X counts as Y in C) can be reconstructed as a regulative rule, by using adequate correlation devices.\(^{67}\) The rule can be rewritten as: “If X lies north of the river (a given correlation device), then it is my property, and if that is my property, then I have the right to use it. If X does not, then it is not, and I have no right over it.” As we have seen in the previous section, this rule can be shown to be a correlated equilibrium of a coordination game (hence, the “rules-in-equilibria” approach).

Searle has two replies to Guala and Hindriks. First, he doubts that this procedure of describing property rights has any real-world equivalent.\(^{68}\) As we have discussed in the previous section, we can have doubts that any real world institutional fact actually corresponds to Guala and Hindriks’s procedure. Do people count something as their property because they have come to an equilibrium with their neighbor that what lies to the North of that stone is A’s property and to the South of the same stone is B’s property? This seems far-fetched. Second, Searle argues that they fail to describe what constitutes property.\(^{69}\) In the example just mentioned, people already know what property is and what rights come with it. They just need to decide on the practical limits of the land they own. They are not defining a new institution. Searle argues that constitutive rules are necessary for a theory of institutions because the creation of institutional facts is the creation of a new fact. Regulative rules regulate already existing activities, but they cannot create new activities. Constitutive rules are thus necessary, and they take the general heuristic form of the imposition of a status Y over an object X, where the status operates at a “higher” level than the object.

I think that Guala and Hindriks are putting too much importance on the linguistic formulation of rules. Though regulative and constitutive rules are of course linguistic rules, the difference between

---


\(^{68}\) Searle, ‘Status Functions and Institutional Facts’, 509.

\(^{69}\) Searle, 510–11.
them is not merely in their linguistic formulation. And it has no necessary connection to a particular formulation or expression. Each theorist reviewed here is putting a lot of effort to come out with nice formulas like “X counts as Y in C” or other more esoteric ones. Yet, the key difference between regulative and constitutive rules is not in their formulation, it is in what they do. Constitutive rules, however, you express them, create a new reality. Regulative rules do not. Thus some people were driving on the right side of the road, I guess, before there was a rule saying that they should. The advent of the rule did not create anything new. Now, a newborn baby may move carved pieces of wood on a board accidentally. She is not playing chess though. She will only do so if she knows and understands that moving a certain piece of wood in a certain direction counts as a particular kind of move in chess.

In that light, the rule “If X lies north of the river (a given correlation device), then it is my property, and if that is my property, then I have the right to use it” appears like a weird mix of regulative and constitutive rules. “If X lies north of the river, then it is my property” is regulative. Substituting “south” with “north” does not change anything about the nature of the institution of property. However, “if that is my property, then I have the right to use it” is constitutive. If one writes instead “if that is my property, then I must pray the lord five times a day”, the word “property” takes a different meaning here. As has been argued before, when making that statement, Hindriks and Guala thus presuppose, rather than assert, what property is.

To conclude, a rule of thumb that will perhaps make clearer the difference between regulative and constitutive rules, and show that this difference is relevant, is the following: “if institution A remains seen as A after rule R has become R’, then R is regulative. If not, it is constitutive.” Clearly, there is nothing intrinsic in the formulation of the rule.

Incentives

Why do people stick to an institution? Why do they obey rules? The answer to this question is yet another reason for discord between Searle and several of his critics. Hindriks and Guala, for instance, argue that the failure to explain why people stick to certain institutions is a central flaw of Searle’s overall project, as well as of all rules-based accounts of institutions. They defend their game-theoretic account of institutions partly on the basis that it secures some place to incentives, and thus can explain why people stick to certain institutions and why some institutions sometimes fail. Smit et al. also resort to a similar argument. On the other side, Searle argues that people have often no incentives to respect the rules set out by institutions, and yet follow them. He also shows that sometimes the incentive may disappear and yet people continue to act according to the rules. Some people keep their promises even if they have no interest of doing so, people pay their taxes even in the absence of fines, drivers respect the speed limits even in the absence of the police, etc. Hence, Searle claims that institutions, through the generation of deontic powers, give people desire-independent reasons for action. For instance, my having a duty to keep my promises is often a sufficient reason for fulfilling it, whether or not I have an interest in doing so, and whether or not others will do so as well.

I do not want to dig into the debate here, for I think there is no need for disagreement on these issues. Both sides have a part of the answer. People sometimes do things because they have an incentive to do so. At times, they also do things even if they have no interests or incentives in

---

70 Hindriks and Guala, ‘Institutions, Rules, and Equilibria’, 462.
doing so. Money is a case in point: most people use dollars because they are born in a society (the US) where people use dollars all the time. They have got the habit of using them, and often do not even think about it.73 Others use dollars because the law forces them to do so (e.g. to pay their taxes). Still others do so out of self-interest. As an example, one could take the case of Argentinians who resorted to use dollars when their own currencies failed, or the more mundane daily operations of multinationals that trade in the currency that best fits their needs. The explanation of behavior does not reduce to one mechanism only, and perhaps there are many more of them than Searle, Guala and Hindriks might think.74

**Conclusion**

This paper has first summarized Searle’s theory of institutions, with the example of money in mind. It then reviewed several debates on particular bits of his theory. I have discussed the (lack of) physical nature of money; the role of functions; the role of collective intentionality; the constitutive/regulative rule distinction; and the role of incentives in the design of specific institutions.

A final point is in order. Many social scientists have been wondering what the point of social ontology might be.75 Many regret that Searle’s approach is too narrow, and that it ignores plenty of social facts that the social sciences consider as essential to their business.76 Searle does not deny it: he makes clear that his sole focus is on institutional facts, and that he leaves aside a lot of what constitutes social reality.77 What, after all, may be wrong with this narrow focus?

What may be wrong is that his theory could be of no use for the social sciences. Searle is actually agnostic on the possible uses of his theory for the social sciences. He writes that “he doesn’t really know” what these could be.78 Many philosophers disagree. They think that social ontology ought to play a role for the social sciences. On one side, Epstein79 claims that social ontology should lay the foundations for proper social sciences. On the other side, Guala80 argues that the role of social ontology is to bridge the gap between philosophy and the social sciences, without any of them taking a superior or foundational position.

Shall these projects succeed? I don’t know. One might worry that Searle’s, Guala’s and Epstein’s attempts are far too abstract and detached from real world problems to be of any help.81 As Sugden82 argues, moreover, economics (and other social sciences) has been doing quite well without social ontology so far, and it is up to the philosophical camp to show that taking it into account would bring significant benefits. After all, economists and social scientists have attempted to explain the emergence, stability and change of institutions well before Searle, Guala and Epstein

---

74 Elster, *Explaining Social Behavior*.
75 Lukes, ‘Searle and His Critics’.
79 Epstein, *The Ant Trap*.
80 Guala, *Understanding Institutions*.
started out their projects, and even afterwards, without noticing their existence. For my part, I share Aydinonat and Ylikoski’s view that what the social sciences need, and what matters to them, is a focus on “particular institutions as historical causal complexes, not on their postulated abstract functional essences.” Social ontology is an immensely interesting field of study, and it will remain so for a long time, but its task is simply not the same as the one of the social sciences. As Searle himself acknowledges, the work of social scientists may well begin where the philosopher’s job ends.

---

84 Greif, *Institutions and the Path to the Modern Economy*.
Bibliography


