Acceptance Is Not Enough, but Texts Alone Achieve Nothing: A Critique of Two Conceptions in Institutional Ontology

CORRADO ROVERSI
University of Bologna – School of Law & CIRSFID
Via Galliera 3
40121 Bologna

corrado.roversi@unibo.it
Tel. +39 51 277203

ABSTRACT
In this paper, I draw on contemporary social ontology to deal with the classic legal-philosophical problem of the ontology of institutions. Institutions are made up of rules—of constitutive rules that define most institutional procedures, roles, and facts. Hence, if we are to investigate the question of institutional ontology, we must conceive it in light of the traditional legal-philosophical problem of the ontology of rules. I thus criticize two competing views developed in social ontology in the effort to pin down the ontology of constitutive rules: the collective-acceptance view (which I treat under the heading of “rule realism”) and the textualist view (which I treat under the heading of “rule positivism”). I argue that although each view captures important aspects of the ontology of institutions, neither captures them all.
Acceptance Is Not Enough, but Texts Alone Achieve Nothing: A Critique of Two Conceptions in Institutional Ontology

1. Introduction: Rule Positivism and Rule Realism

Social ontology is an active and growing field of interest for contemporary philosophers. Apart from the seminal work done by authors like John Searle, Raimo Tuomela, Margaret Gilbert, and Seumas Miller, there is much multidisciplinary research (some of it conducted within the European Network of Social Ontology, or ENSO) in analytic philosophy, developmental psychology, and economics, as well as in information science, which is achieving significant results by assuming social ontology as a common framework. This framework is making it possible for scholars from different backgrounds to explore, among many other problems, the nature of social action, the role and limits of intention in such action, the development of social reality in children’s perceptions, and the simulation of social interaction through artificial intelligence.

All these studies come to bear directly on the core problem of institutional ontology, that of understanding the peculiar nature of institutional reality. And it is striking, in light of this fact, social ontology hasn’t yet become a major field of interest for contemporary legal philosophers, too. Current work in social ontology is not having much impact on the discussion about the classical ontological problem of legal philosophy, that of explaining the reality of law and legal institutions. And, accordingly, unlike what is going on in the theoretical and experimental areas of investigation just mentioned, legal philosophy is making little use of the basic concepts of social ontology, concepts such as those of constitutive rule, collective intentionality, institutional fact, meta-institutional terms, status function, and practical import.

But in the general picture developed in social ontology, institutions are in the first instance made up of rules, and in particular (in Searle’s terminology) of constitutive rules: these rules define most institutional procedures, roles, and facts. And so, if we are to thrash out the question of institutional ontology, we cannot but place the discussion within a broader discourse, that of the ontology of rules in general, and this does figure as a traditional problem in legal philosophy. In fact, apart from the obvious reference to Wittgenstein’s Philosophical Investigations and to the ensuing and seemingly never-ending rule-following debate (a debate mainly concerned with the nature of semantic normativity), the problem of figuring out the ontological nature of rules has always been a central theoretical concern of legal philosophy, one that that has been debated for centuries with regard to legal rules and norms.

I will thus discuss this classic legal-philosophical problem by considering how it has been addressed in social ontology. A distinction can be drawn in this area of study between two competing views of the ontology of constitutive rules: I will call these views rule realism and rule positivism. On a rule-realist view, the rules constitutive of an institution can exist only as part of the causal (mental or behavioural) process through which the institutional activity they constitute is practiced. These rules must accordingly be conceived as something existing as part of the institutional activity’s practice. On a rule-positivist view, by contrast, the rules constitutive of an institution can exist before and independently of the causal process through which the institutional activity they constitute is practiced. On this view, these rules’ existence is a separate matter from that of their effectiveness (in the legal-
philosophical sense of this term), for their independent status endows them with an
objectivity that cannot be traced to the behaviour or the mental states of those who are
engaged in the institutional practice constituted by the same rules.¹

I should now briefly elaborate on what I mean by “realism” and “positivism” in connection
with the idea of a rule. As I suggested, these are broad headings we are dealing with—we
might call them archetypes—each of them covering a whole cluster of conceptions in
contemporary social ontology. But, as was also suggested, these two archetypes, dealing
with the ontology of rules in general, have some clear points in common with those
conceptions in legal philosophy that take as their object the ontological nature of legal
norms. Now, as the reader may know, writers on legal philosophy who believe that the
existence of a legal norm cannot be separated from its effectiveness—and hence from the
behaviour and/or mental states of those who follow that norm—qualify themselves as legal
realists. The label “realism” thus suggests itself as the natural choice in describing this view,
and I have replaced the qualifier “legal” with “rule” only to make it clear that we are
considering this question from a broader perspective, that of social ontology, which is
concerned with the rules constitutive of institutions in general, rather than with legal
institutions in particular.² Less straightforward is the label “rule positivism.” For in this
case the qualifier “positivism” should be understood not so much in connection with the
traditional legal-philosophical conception commonly called “legal positivism” as in
connection with a broad concept of rules considered as objective, that is, as posited and
hence existing, independently of what we think about those rules and how we relate to
them in action. In fact the relation between rule positivism and legal positivism would be a
subject in its own right, complex enough to take up an entire book. But at least this much
can be said, which is that there is a basic categorial difference between the two
conceptions, the former (rule positivism) dealing with constitutive rules and the latter (legal
positivism) with legal ones: two different kinds of rules are involved, belonging to two
different taxonomies, as can be appreciated from the fact that it makes perfect sense to
speak without redundancy of legal constitutive rules. I should also stress that, unlike legal
positivism, rule positivism need not call into play a legal or institutional enacting authority
as the source of the rules in question: what is essential (on a rule-positivist view) is not a
rule’s connection with authority but its objectivity, or rather, our ability to find
“independent criteria” for determining the existence of social facts and entities. Hence, rule
positivism could aptly be stated in the negative: An institution’s constitutive rules cannot
exist only as mental states or as patterns implicit in the behaviour they constitute. Their
existence must be objective, that is, independent; or, stated otherwise, they must depend
for their existence on something other than the mental states of those who follow them or
the behaviour through which they can be observed.

¹ Of course, I am not claiming the distinction between rule realism and rule positivism to be
exhaustive (there are other ways of conceiving constitutive rules), but I do think this distinction
gives us a good vantage point from which to address a crucial question that constitutive rules
specifically raise: that of their relation to the object they constitute.

² It must be stressed here, once more, that rule realism is being presented as an archetype, rather
than as a historico-philosophical account of legal realism, for which reason it cannot reflect the
nuance one finds in the work of authors like Axel Hägerström, Karl Olivecrona, Alf Ross, Oliver
W. Holmes, and Leon Petrażycki, among many others.
In the two sections that follow, I will argue that rule realism and rule positivism both fall short, each in its own way, as ontologies for constitutive rules and hence as conceptions of institutional ontology. In Section 2, I show that rule realism makes a mistake in equating the existence of constitutive rules with their recognition, for in so doing it fails to account for the objectivity of constitutive rules, an objectivity which is what instead lies at the core of the rule-positivist view. Conversely, in Section 3, I maintain that rule positivism, for its part, in its quest for objectivity, too readily dismisses the fact that constitutive rules must be recognized if they are to have an effective ontological bearing on institutional reality, and this element of recognition is exactly what rule realism singles out as a fundamental condition for a rule to exist. My conclusion will thus be that rule realism and rule positivism each capture important aspects of the social-ontological problem we are concerned with, but that neither conception can take on the problem in full breadth.

2. Acceptance Is Not Enough ...

It is certainly a compelling argument that must be put forward in attempting any criticism of rule realism, considering that this is without question the dominant theory in social ontology. Indeed, the vast majority of those who have taken up the question of the reality of institutions in social ontology have treated the matter by invoking what for ease of reference we might call “collective acceptance,” the idea being that an institution, institutional fact, or constitutive rule exists only insofar as this institution, fact, or rule forms the subject of collective acceptance. On this conception, then, an institution’s reality is equated with its life “on the ground”: its reality is made to depend on the existence of a shared or at least a widespread practice of acceptance.

When it comes to spelling out what collective acceptance means, however, the theories put forward by writers on social ontology begin to diverge. According to Searle, for example, institutional facts require an attribution of status functions—namely, functions that can be performed independently of physical features—and this attribution can be made only by means of collective acceptance; collective acceptance, in turn, is a kind of collective intentionality, an intentionality “in the plural” that can be considered a primitive feature of human social action. Hence, all kinds of social action can be analyzed in terms of given groups of individuals having mental or intentional states with the content “we intend that $x$, and institutional facts can be seen as a particular case of this general structure, a case in which those who make up a given community have intentional states with the content “we accept that ($\alpha$ has a status function, and thus has deontic powers).” In Searle’s theory, constitutive rules are a particular case in which a status function is attributed to something, and this means that not all such cases can be described in terms of constitutive rules (or “standing declarations,” as Searle has latterly been expressing himself): we can speak of constitutive rules only in those cases that can be analyzed in terms of the individuals making up a given community having intentional states with the content “we accept that (constitutive rule).” This theory falls squarely within rule realism, since constitutive rules exist

---

here only as instances of collective social actions and must be traced to the mental states accompanying the activity these rules constitute.

It is not only Searle, however, who espouses rule realism. There is at least one important rule-realist alternative to Searle, a theory that accounts for social-institutional reality in terms of mutual beliefs, or “common knowledge” (as philosophers sometimes say). To be sure, this too is more a grouping of theories than a single theory, since the specifics—those from which flow important consequences—are almost invariably worked out in different ways in the different accounts of institutional reality in terms of mutual beliefs (with authors such as Margareth Gilbert, Michael Bratman, Raimo Tuomela, Eerik Lagerspetz, and Seumas Miller). But to simplify matters, I will consider these accounts under a single heading, a theoretical archetype, so to speak: a model or outline that all those who have worked in this direction would take up as their basic frame of reference. On this (archetypical) theory, institutional facts exist only if they are collectively recognized or accepted to exist, and this collective recognition or acceptance can be analyzed in terms of individual mental or intentional states plus mutual beliefs, that is, recursive beliefs about others’ beliefs. It we apply this archetype to constitutive rules (the building blocks of institutional facts), we will find that this, too, clearly amounts to a form of rule realism, because constitutive rules exist only as instances of collective mental states embedded in the activity these rules constitute.

The argument I will be presenting criticizes rule realism as a broad conception, regardless of whether collective acceptance is construed as a mental phenomenon (a collective mental fact) or as something to be inferred from the way people generally act (from a practice). Which means that I intend the argument to apply to behavioural and mentalistic versions of rule realism alike. Likewise, if this argument holds up, it will do so independently of the view one takes on three contentious issues arising in connection with the collective-acceptance thesis. The argument will not be affected by the way we work out (i) the question of how to describe the “structure” of the intentional states that make up collective acceptance, and in particular the question of whether these states are to be conceived as individual intentions plus mutual beliefs or as intentions in the plural form “we”;


intentionality or a social one; or (iii) the question whether collective acceptance exists only as a belief (merely a factual matter of what is true or false), as active participation, or as passive recognition. The argument instead proceeds from a threefold analytical distinction I should like to bring to bear on the view that institutional reality can be reduced to the existence of intentional or mental states in the minds of those who recognize the relevant institutions. I would say that once we take this view, we should be sufficiently clear about whether these states are conscious and whether they are supposed to be under the conscious control of the individuals concerned. We thus ask: Are individuals aware of these mental states? And, if so, is it within these individuals’ power to exercise at least some degree of control in forming these states and fixing their content? From this we get three possible ways in which an institution’s constitutive rules can be thought to (ontologically) “lie in” a mental or intentional state: Constitutive rules can exist (a) as conscious mental or intentional states whose formation and content individuals can control (I will call this the individual-control case); or (b) as tacit mental or intentional states whose formation and content individuals cannot control (the underground case); or (c) as conscious mental or intentional states whose formation and content individuals cannot control (the collective-control case). I take it there are no other practical possibilities, aside from these three cases, in mapping out the question of whether individuals are conscious of, and can exercise control over, the mental states credited with creating an institutional reality. I will treat these three cases separately and show that in none of them can an institution’s constitutive rules be thought to exist just because they form part of the content of mental or intentional states.

**Step 1: Rebutting Rule Realism in the Individual-Control Case.** As we just saw, on the individual-control view, an institution’s constitutive rules exist solely by virtue of their forming part of the content of mental or intentional states whose formation and content individuals can control. I maintain that this understanding of rule realism fails in two ways: it turns out to

---

6 Searle has famously held that collective intentionality is to be understood as part of our biological makeup, something that “could be had by a brain in a vat.” See, for example, J. R. Searle, ‘Collective Intentions and Actions’, in *Consciousness and Language* (Cambridge: Cambridge University Press, 2002), p. 90–105 (see esp. p. 97). This internalism with regard to collective intentionality has been firmly criticized by authors who have insisted on the social nature of collective intentionality. See, for example, D. Fitzpatrick, ‘The Self-Defeating Nature of Internalism with Respect to Social Facts’, *American Journal of Economics and Sociology*, 62 (2003), pp. 45–66; and A. W. M. Meijers, ‘Can Collective Intentionality Be Individualized?’, *American Journal of Economics and Sociology*, 62 (2003), pp. 168–83.

7 Eerik Lagerspetz and Bruno Celano, among others, have criticized the view that social institutions can be made to depend on beliefs. See, for example, E. Lagerspetz, ‘Institutional Facts, Performativity and False Beliefs’, *Cognitive Systems Research*, 7 (2006), pp. 298–306; and B. Celano, ‘Intenzionalitá collettiva, false credenze’, in *Ontologia sociale e regole costitutive*, ed. by P. Di Lucia (Macerata: Quodlibet, 2003), pp. 71–98. But it is probably Tuomela who has devoted the greatest attention to the difference between collective acceptance as a kind of belief and collective acceptance as a kind of intention. See, for example, Tuomela, *The Philosophy of Social Practices*, pp. 151ff.

8 A careful reader will have noticed that this threefold distinction is based on two variables—namely, whether a mental or intentional state is conscious or tacit, and whether or not we can control its formation and content—from which follows a fourth case, that of tacit mental or intentional states whose formation and content individuals can control. The reason I do not discuss this fourth case is precisely that I understand it to be impossible.
be conceptually inconsistent, and it fails to reflect the way institutional phenomena actually work.

Let me first take up the conceptual inconsistency. The basic idea is that it is impossible for rule realism to hold itself out as a thesis about rules (about an institution’s constitutive rules) while construing rules as existing only as part of mental or intentional states whose formation and content individuals can control, for there is no way that we could call those entities “rules.” Indeed, if all it takes for an institution’s constitutive rules to exist is a mental state in which they are “housed,” a place in which their content is shaped, so that the individual to whom we ascribe that mental state can change them at will, then there would be no way for these rules to have any diachronic or synchronic stability, fixedness, or identity: we could each individually change their content (just by changing our minds as to what the rule is), and different individuals or groups could easily have different understandings of the same rule at the same time, without even recognizing it as the same rule, which means that the same rule is really a bunch of different rules. Clearly, such wildly swinging rules could not in any way guide our conduct in any ordinary setting, much less in an institutional one, because a rule (especially a practical rule) can be such only if it can serve as guidance for our behaviour, and this cannot happen privately: there needs to be a public space for a rule to work, and that goes double for a constitutive rule in an institutional setting. This is a distinctively Wittgensteinian point, which I believe can be brought to bear on constitutive rules:

[T]o think one is obeying a rule is not to obey a rule. Hence it is not possible to obey a rule “privately”: otherwise thinking one was obeying a rule would be the same thing as obeying it.  

I take this to be a statement on the concept of a rule. As Wittgenstein clearly states:

Is what we call “obeying a rule” something that it would be possible for only one man to do, and to do only once in his life?—This is of course a note on the grammar of the expression “to obey a rule.”

Now, as I suggested, if this point holds for rules in general, it does all the more so for those constitutive rules through which institutions are created. Wittgenstein says that “to obey a rule, to make a report, [...] to play a game of chess are customs (uses, institutions),” this precisely to contrast institutions with the private dimension, and in fact something like an “individual institution” would have to amount to a contradictio in adiecto. It follows that the very concept of an institution is inconsistent with individual control, and rule realism so

---

10 Ibid., § 199.
11 Ibid.
interpreted is accordingly inconsistent with the concept of a rule constitutive of institutions.

Indeed, it is quite clear that institutional phenomena are not the outcome of individual deliberation, for otherwise they could not in any way influence anyone’s behaviour—something they instead ordinarily do under appropriate circumstances. Consider the example of a simple institution such as a game. Clearly, individual players cannot just decide on their own what the rules of the game are, unless we presuppose that they have to engage one another in the game on the basis of a “mutually preestablished harmony,” which at the very least is an unlikely scenario. In fact, if each player was free to make up the rules of the game while playing, there would be no difference between cheating and playing, and any player’s behaviour could count as institutionally relevant so long as this player had a rule on which basis to explain such behaviour. But this is not how institutions work. Social rituals and norms are so called because they take root among groups: they accordingly tend to be more “entrenched” in such groups, and this entrenchment excludes in no uncertain terms that these rituals and norms can depend on individual deliberation. The individual-control view of rule realism therefore openly contradicts institutional reality: if we take the view that institutions are created through the content of individually controlled mental states, then we would be hard put to it to explain the necessary fixedness that institutions need to attain if they are to be functional, a fixedness that in turn requires a shared, public setting or framework—precisely what individual control cannot secure.

Two points are worth making before we move on to other, more promising versions of rule realism. First, a rebuttal of the individual-control case does not entail a blanket dismissal of all individualist accounts of institutional activities. This also holds for the most resolutely individualist approaches to the ontology of social and institutional practices. I am thinking in particular of the theory that Stephen Turner expounds in his two books *The Social Theory of Practices* and *Explaining the Normative,* where he argues that social and institutional practices cannot be accounted for in terms of shared objects, whether these are understood as shared causes for individual institutional actions, or as shared presuppositions, or as shared norms. Turner rejects all those approaches that hypostatize practices, norms, and institutions as supraindividual objects (among these approaches he includes Searle’s theory of collective intentionality), and he advocates an individualist view that instead explains institutional and social activities in terms of individual habits and dispositions. Let it be noted that such an individualist view, whatever merits it may have, cannot in any case be traced to the individual-control variety of rule realism. In fact, even Turner’s flat-out individualist account of social reality must take into account an element capable of assuring a certain degree of objectivity, such that individuals do not have complete control over institutional behaviour. This element, in Turner’s account, is individual innate empathy rooted in mirror neurons plus the necessary feedback coming from others, and these factors are essentially intersubjective. Indeed, it is Turner’s view that “empathic projections” must be “confirmed, sustained, corrected and improved through interaction with others.”

---


14 Ibid., p. 178.
out to intersubjectivity—which is exactly the opposite outcome of what rule realism arrives at in the individual-control case.

The second point I should like to make is that the individual-control account of rule realism also highlights a problem that Searle’s so-called internalism comes up against in explaining collective intentionality, a problem arising from his well-known claim that we could have collective intentionality even if we were “brains in a vat” (see footnote 6). Clearly, this cannot be taken to mean that individuals have complete control over the formation and content of their states of collective intentionality, for otherwise we would lapse into the individual-control case, which for the reasons previously outlined is unacceptable. But if individuals do not have complete control of their collective intentional states, yet at the same time can form such states in complete isolation from others, then the question arises as to what the formation and content of such states depends on: on what basis, for example, can these states be deemed appropriate? Indeed, it is quite clear that we must suppose there to exist some kind of external, social control in determining the formation and content of the individual collective intentional states thought to lie at the root of institutions. But then, in this case, Searle must explain how his theory of collective intentionality accounts for this social dimension. I am not sure that, even in his recent book, Making the Social World, he has managed to do so.

Step 2: Rebutting Rule Realism in the Underground Case. On the account of rule realism I have called the underground case, an institution’s constitutive rules exist only as part of the content of tacit mental or intentional states whose formation and content individuals cannot control. I consider this a behavioural version of rule realism, because on this account we cannot analyze these underground mental states if not by looking at the behaviour individuals actually have, which means we have to derive the content of those states ex post from what actually happens in practice. The rule reveals itself, or becomes manifest, in the institutional practice.

Underground rule realism has been advocated by scholars who have treated social practices from a Wittgensteinian point of view. Theodore R. Schatzki, for example, has argued that the “we” standing behind social practices is nothing more than a matter of behavioural mutual intelligibility:

> In most cases [...] a person eventually becomes “one of us.” This means, first of all, that he or she reacts verbally as we do [...] Wittgenstein’s way of putting this is that the person masters the technique of language, where “technique” means customary and regularized way of using or, more perspicuously, customary and regularized way of behaving [...]. Being one of us means, more broadly, that a person speaks and behaves intelligibly to us.16

At first sight, this leads to a conception that seems different to that put forward by Searle, Tuomela, and the other collective-intentionality theorists. On this conception, we cannot

---


explain social practices by appealing to a mental state that \((a)\) is shared among individuals and \((b)\) is causally responsible for an individual's actions (as in the case of Searle and Tuomela's we-intentions); rather, we must appeal to behavioural evidence in terms of mutual understanding. A similar explanation of social actions has been provided by David Bloor, with special attention to social institutions. According to Bloor, institutions exist essentially as a matter of behavioural consistency among different persons, in just the same way as happens with social practices in Schatzki's view. He takes money as an example:

> Metal discs are coins because they are called “coins.” My references to things being called “coins” is shorthand for the entire repertoire of behaviour associated with their being thought of, or regarded as, or treated as, coins [...]. On this usage, speaking of a thing as a coin isn’t meant to refer to a purely verbal act, but to the whole pattern of behaviour into which such explicit verbalisations are woven. \(^{17}\)

As is the case with Schatzki’s linguistic social practices, so in this reconstruction by Bloor the question whether an institution exists, how it comes to exist, and “where” it exists, is something we can only work out \textit{ex post} by observing the participants’ behaviour and recognizing that they understand one another. This is a \textit{regressive} explanation of social reality, that is, an explanation that starts out from observed practice and from it derives an institution’s existence, whereas the collective-acceptance view typically offers a \textit{progressive} explanation, one that first postulates an institution’s existence in the form of mental entities and then proceeds to explain how the mentally lodged institution can bring about a corresponding practice. On the regressive explanation, the institution \textit{is} the corresponding practice, while on the progressive one, it is what \textit{causes} the practice. This distinction, however, becomes quite blurred if mental states and current practice are made to coincide, as when mental states are understood as \textit{tacit} entities, that is, as behavioural dispositions that can reveal themselves only in action. And this conception of collective intentionality as a tacit phenomenon is not at all rejected by authors such as Searle and Tuomela. Searle’s concept of \textit{background} is meant precisely to describe this dimension, and an analogous insistence on the possibility of tacit we-intentions can be detected in Tuomela. \(^{18}\)

A typical feature of behavioural regressive explanation, where an institution’s existence is always explained \textit{ex post} (that is, starting from a practice), is that the content of an institution's constitutive rules shows through in particular when something in that practice “goes wrong”—or, as Tuomela says, “in cases of institutional breakdown.” \(^{19}\) It is quite striking that this kind of explanation should lie at the core of the method used by John Langshaw Austin to explain the structure of illocutionary acts, and hence was central to


that institutionalization of language which for several years has been the hallmark of Searle’s work on speech acts. Consider this methodological remark by Austin:

Besides the uttering of the words of the so-called performative, a good many other things have as a general rule to be right and to go right if we are to be said to have happily brought off our action. What these are we may hope to discover by looking at and classifying types of cases in which something goes wrong and the act—marrying, betting, bequeathing, christening, or what not—is therefore at least to some extent a failure: the utterance is then, we may say, not indeed false but in general unhappy. And for this reason we call the doctrine of the things that can be and go wrong on the occasion of such utterances, the doctrine of the *Infelicities*.20

And observe now how Searle applies this regressive method to the speech act of promising. He starts out from the Austinian “infelicities” and proceeds to the constitutive conditions of performatives:

A promise involves an expression of intention, whether sincere or insincere. So to allow for insincere promises, we need only to revise our conditions to state that the speaker takes responsibility for having the intention rather than stating that he actually has it. A clue that the speaker does take such responsibility is the fact that he *could not say without absurdity*, e.g., “I promise to do A but I do not intend to do A.”21

Now, if we consider the peculiar status of language conceived as an institution, coupled with the regressive explanation attempted here by Austin and Searle, we can see how rule realism turns conceptually inconsistent in the underground case to the extent that it relies on a regressive explanation of constitutive rules. To see this, we can start from an observation that Paul Amselek makes about language and constitutive rules:

The analogy between the rules constitutive of certain kinds of games and the rules constitutive of certain kinds of speech acts is in effect misleading. [...] With football and other games, essence [at the outset] preceded existence, that is, the game was first conceived and invented—and the rules that constitute it were created and circulated—and then the game so conceived was practiced. On the contrary, with speech acts (at least the most fundamental ones), it was existence [at the outset] that preceded essence. [...] In other words, for these fundamental speech acts a sort of *ex post* institutionalization was effected, an institutionalization which gives rise to an


optical illusion and makes us believe that the rules formulated *ex post* to define these kinds of speech acts precede the corresponding practice. But that is wrong.\(^{22}\)

The point here is clearly stated. It would be quite a stretch to construe the fundamental speech acts as dependent on rules, and all the more so when dealing with *constitutive* rules, that is, rules that can be such only if they precede and make possible the practice they constitute. For this reason, we can take this explanation of linguistic practices to be regressive, as Austin and Searle seem to do. But what we obtain by so doing are not in a proper sense constitutive rules but *ex post* explanations revealing how the practices in question are structured as a matter of fact: instead of laying out what makes a practice possible to begin with (the proper role and function of constitutive rules), we describe the practice such as it already exists. And so, in the outcome, we get not constitutive rules but what might be called “statements on constitution,” which are rather descriptions than declarations: they do not constitute but *describe* institutional structures, and as such they can even be true or false (a property we could not ascribe to constitutive rules).

Constitutive rules can in fact be confused with statements on constitution because they appear to be descriptive. Take, for example, Searle’s formula of constitutive rules, “X counts as Y in context C.” There is nothing distinctively regulative about this formula, and so it can lead one to think that constitutive rules are in some sense descriptive. But, as Searle has pointed out on many occasions, constitutive rules do something other than describe: their “direction of fit” is not simply word-to-world, as in the case of simple descriptions, but rather two-way—at once word-to-world and world-to-word—which is to say that constitutive rules *describe* the very reality they *constitute*. Statements on constitution, by contrast, describe something that does not depend on them, a practice whose concept and performance is not made possible by those statements. In *Making the Social World*, Searle himself implicitly recognizes this difference between *constituting*, on the one hand, and *describing* a constitution, on the other:

At first sight, it might seem that formulae of the form “X counts as Y in C” function the same for language as they do for other institutional facts. [...] But in spite of this apparent similarity, there is a huge difference, and it has to do with the nature of meaning. [...] In the case of the sentence, formulae of the form “X counts as Y in C” describe the constitution of meaning and not a separate linguistic operation that we perform. But in the case of nonlinguistic institutional facts, constitutive rules of the form “X counts as Y in C” describe a linguistic operation that we perform by which we create new institutional facts [...]\(^{23}\)


\(^{23}\) J. R. Searle, *Making the Social World: The Structure of Human Civilization* (Oxford: Oxford University Press, 2010), pp. 14–5; italics added; cf. pp. 111–2. Searle makes this statement in rejecting the institutional theory of language he had advocated since the 1960s, a theory that would ultimately lead him to circularity. The charge of circularity was prompted by this question: If institutions depend on language to exist, how can language be an institution in its own turn? Searle has been conscious of this problem at least since *The Construction of Social Reality*, where he rejects constitutive
Here, Searle draws a distinction between descriptive “formulae of the form ‘X counts as Y in C,’” on the one hand, and constitutive rules, on the other, a distinction not unlike the one made here between statements on constitution and constitutive rules. It should be clear from the passage just quoted that a regressive explanation, such as the one Searle discusses here with regard to language, is incompatible with the concept of a constitutive rule. A regressive explanation treats constitutive rules as descriptive statements on constitution, but in so doing it empties the very concept of a constitutive rule of any meaning, because statements on constitution do not create any practice: they instead start from an existing practice and describe it (with “formulae,” as Searle says).

In discussing behavioural rule realism, the example of language can be generalized to apply to the whole of institutional life. Rule realism in its “underground” inflection pins the existence of every institution to the existence of an institutional practice, thus taking the practice as its explanatory point of departure. But, again, on such an approach there is no way to distinguish an institution’s purportedly constitutive rules from descriptive statements about the institutional practice’s internal constitution. Thus, behavioural rule realism cannot account for constitutive rules as creative of an institutional practice, and so, consequently, it cannot explain these rules’ ontological peculiarity, either.

This conclusion is interestingly and strikingly complementary to the criticism I previously made of the individual-control statement of rule realism. Just as rule realism was shown in that case to be inconsistent for its failure to explain how constitutive rules can be genuine rules, the brand of rule realism we are considering in the underground case turns out to be inconsistent for its failure to explain how constitutive rules can be constitutive. It is particularly important to point out this complementarity because I am not saying that if we start out in our explanation from an existing practice—thus adopting a regressive argument—we thereby end up denying the normative, rule-bound character of that practice. This is the fault that Robert Brandom finds in the view he calls “regularism,” on which the norms implicit in a practice are reduced to regularities amenable to description. According to Brandom,

The simple regularity approach is committed to identifying the distinction between correct and incorrect performance with that between regular and irregular performance. [...] The progress promised by such a regularity account of proprieties of practice lies in the possibility of specifying the pattern or regularity in purely descriptive terms and then allowing the relation between regular and irregular performance to stand in for the normative distinction between what is correct and what is not.  

As Brandom shows (following in the footsteps of Wittgenstein and Kripke), there is no way to single something out as a regular performance without thereby singling it out the rules as a means to explain language. See J. R. Searle, The Construction of Social Reality (London: Penguin, 1996 (1995)), pp. 72ff.

correct performance. Hence, regularism “loses the contrast between correct and mistaken performance that is of the essence of the sort of normative assessment being reconstructed.” However, Brandom also argues that regularism is not implied by practice-centered theories (behavioural rule realism being one of them), because we can regard norms as implicit in the know-how a practice consists in, thus fashioning the status of correctness into a sort of practical attitude.

However, the option that Brandom leaves open is not something we can apply to constitutive rules, because these rules are supposed to make an institutional practice possible by bringing it into existence or “creating” it. Of course, to this end they must in some sense precede the practice, for otherwise we could not distinguish them from a simple normative practice, and if we take constitutive rules to exist only as normative practices—as behavioural rule realism ultimately does—then it is the existence of the practice that entails the existence of the rules, and not the other way around. Hence, with Brandom, we can allow for the existence of normative practices. But if we are to allow for an institution’s constitutive rules, these rules cannot exist simply as embedded in a practice, for that would make it impossible to distinguish the case in which the practice gives rise to rules from the inverse case in which the rules are constitutive of the practice.

Moreover—apart from these conceptual considerations—rule realism on the underground interpretation seems to contradict the phenomenology of institutions, even in the simplest cases. Consider, again, the example of games. Suppose we have a game with distinct authors, namely, a well-defined group of persons who are aware of their creative role and thus formalize the game very quickly. Our intuitive understanding of this example would preclude the possibility that the rules of the game exist only as implicit in the players’ behaviour, because we know that unless somebody invented that game, setting down its constitutive rules, the behaviour consisting in playing the game could not have happened to begin with. Hence, those rules in some sense existed before the shared practice could take hold among the community of players. Indeed, without those rules, without the rulebook in which they were set down, without the game designers’ game-developing activity, and without the production of a game-publishing house, no community of players could even exist. In this case, it is evident that the established practice of a community of persons practicing an institution—and the very definition of a community whose practice can be relevant for the institution—is ontologically dependent on the rules, and not the other way around.

One, however, could rebut that this view is heavily skewed. “Suppose,” as the objection goes, “that the whole community of players arrives at an accepted or settled interpretation of the game they all play, and that this interpretation corresponds to their practice. What, in this case, would be the content of the rules of the game? Clearly, that which is established by practice; and this is where behavioural rule realism shows its strength, namely, in providing a ‘home’ for institutional rules and facts, a place where they can be said to exist and where they can be observed and even measured, up to a point.”

I think this objection is off the mark. That an existing practice can be observed does not necessarily mean that this is where we should look in identifying an institution’s

---

25 Ibid., p. 29.
26 Ibid., pp. 25, 30–46.
constitutive rules, especially when that institution has an identifiable author. Where institutional phenomena depend on an explicit act of creation, it becomes conceptually fundamental to refer to that founding act in judging whether an institutional fact or event of a given kind subsists or has taken place. Suppose our game’s original developers issued a statement declaring the community’s shared interpretation of the rules to be wrong. Would such a statement be inconsequential? I think it wouldn’t. For in that event, the community would be forced to rethink its practice, or at least a debate would arise on the merits of the original plan as against current practice. It may happen that the game is tweaked with changes designed to reflect current practice, but this would mean that the community has moved to a modified version of the game: the original game is no longer being played. Or, similarly, the community of players could say, “Okay, the game we’re playing is not the game we thought we were playing, but that’s just as good: Let’s keep playing it anyway!” Any such statement would mean that the players see their current practice as defining a new game, a situation they would probably want to formalize soon enough with a new name for the community’s game and a new set of rules, thus recognizing that the whole community has performed a new act of institutional creation, albeit only in an informal, inexplicit manner.

Now consider the case of this new game that has emerged from a standing practice. Here, a group of players takes up new rules and thus creates a new institutional vocabulary without quite being aware of its own creative role, at least at the outset. But even in this case there will be a point at which the institution will be considered to exist and be established, and from that point on the broad patterns of behaviour within the community (the behavioural “regularities”) will have to be checked against the game’s original statement so as to settle questions of correctness. And this shows that, even in this case, a broad behavioural convergence within the community (its established practice) will not indefinitely be held up as the last word on the game. Once the game as practiced by the community settles into an established and hence recognizable form, it will be fundamental to refer to this established form to decide questions regarding correctness or incorrectness in institutional behaviour.

The conclusion is that rule realism in the underground case is flawed on both conceptual and phenomenological grounds, for on the one hand it misconceives the role of constitutive rules (ascribing to them a role that cannot in fact be constitutive), and on the other it overstates the bearing a community practice can have on an institution’s existence. Let us then see whether rule realism in the collective-control case can fix these problems.

Step 3: Rebutting Rule Realism in the Collective-Control Case. On the collective-control understanding of rule realism, constitutive rules and institutional facts exist only as part of explicit mental or intentional states whose formation and content individuals cannot fully control. The reason for this lack of control is that the formation and content of those mental states is at least in part a collective or social phenomenon, and so I take this variety of rule realism to be roughly equivalent to the collective-acceptance view as considered at the beginning of this section. In short, the question as concerns rule realism in the collective-control case is the following: Are an institution’s constitutive rules belief-dependent? Do they essentially exist in our collective beliefs, and is it these beliefs that shape their content? I believe not, and I explain why by breaking this thesis down into two subtheses. First, I will consider whether a collective belief is always a sufficient condition for a constitutive rule to exist: this is the question of whether community consensus can create a constitutive rule in the community in question. Second, I will consider whether a collective belief is a necessary condition for a constitutive rule to exist: this is the question of whether a constitutive rule can exist even without a corresponding collective belief within the community in question.

Let us take up the first question by supposing that the answer is yes, collective beliefs do make a sufficient condition for the existence of an institution’s constitutive rules, or, more
formally stated, there is a definite set of agents $a_1, \ldots, a_k$ who “share” some mental or intentional states that in turn determine the content of a given institution $I$. And now we ask: Does the community to which $I$ is relevant necessarily comprise only agents $a_1, \ldots, a_k$? The answer is no even on the collective-acceptance view, because collective-acceptance theorists concede that it is perfectly possible for at least one agent $a_w$ not included in the set $a_1, \ldots, a_k$ to be mistaken about the content of the institution $I$ while still being part of the community to which $I$ is relevant. Hence, a single agent’s error does not threaten the collective-acceptance view, because insofar as a sufficient number of agents share the relevant collective intentional states about $I$, it is possible for these agents to correct any individual $a_w$ on the content of $I$. But, clearly, there is a fundamental difference, on this theory, between individual and collective error about the content of an institution’s constitutive rules: individual error can be corrected by reference to the content of the community’s collective intentional states (let us suppose that this content is the same for all individuals and that it can be easily stated); collective error, on the contrary, cannot be corrected, because a community’s collective intentional states are precisely what on this view determines the content of institutions, such that to correct the collective intentional state is to change the institution itself.27

Now, let us assume that on the collective-control account of rule realism, the collective intentional states constitutive of a given institution are conscious states, that is, agents $a_1, \ldots, a_k$ are conscious of these states’ content and constitutive role. My question, then, is: What if instead of single agent $a_w$ disagreeing with $a_1, \ldots, a_k$ on the content of a given constitutive rule, the disagreement was with a definite set of agents $a_w, \ldots, a_z$, a group comparable in size and relevance to the original one? Clearly, what we would be looking at in this case is a substantial divergence within a community on how a given institution is to be interpreted. But if we are to take rule realism in earnest on this interpretation, that is, if we are to believe that all the agents involved in the debate are actually conscious of the constitutive role their intentional states have on the content of their institutions, the dispute could be settled just by looking at the content of each agent’s beliefs. On this reading, however, the game of institutional debate would become a game essentially concerned with the participants’ expressions of belief, that is, in terms of speech act theory, it would be an expressive language game rather than an assertive one. And clearly this is not an accurate description of the institutional domain. Thus, for example, when a jurist says what the content of a given legal norm is, the point is not to express what he or she or a given community thinks but rather to say something true or false about the content of a given norm, something by which to justify their beliefs. And this goes for all kinds of institutional phenomena, from the game of chess to Catholic ritual. Institutional discourse is an assertive language game, its point being to determine the real content of the institution, not to compare expressions of belief.

The collective-acceptance theory could here retreat into a more plausible account, that is, it could argue that agents $a_1, \ldots, a_z$ (all the agents involved in the debate) need not be aware of the constitutive role their beliefs play. Agents $a_1, \ldots, a_z$ can be clear in their minds about the content of constitutive rules, all the while failing to realize that the rules and their content play a role in constituting institution $I$. This means that the agents are treating the institution as something independent of their collective beliefs, and the relative institutional

discourse as assertive, even though that is not the case. Some passages by Searle seem to suggest this view. Consider, for example, the following:

To what extent can people be mistaken about the Y status function? So, for example, people can believe that it is a marriage only if it is made in heaven, and yet on my account, it is still a marriage even if it is not made in heaven. The crucial question is, What rights and obligations does the couple have? How do people actually treat them and how do they regard themselves? If the members of the community accord marital status to them, then whether they do that on the basis of false beliefs is really irrelevant. [...] The acceptance of an institutional fact, or indeed, of a whole system of status functions, may be based on false beliefs. From the point of institutional analysis, it does not matter whether the beliefs are true or false. It only matters whether the people do in fact collectively recognize or accept the system of status functions.28

But then I would have to ask the collective-acceptance theorist: If the agents themselves do not realize their intentional states constitute the content of an institution, how can you? What is the external vantage point from which you can judge that the whole institutional discourse, which seems to be assertive, is only apparently so? How can you rule out the possibility that those intentional states, which you assume agents to share, are not constitutive of the institution but simply mirror the agents’ conception of that institution? This conception may very well be descriptive and not constitutive: it might be a conception built by agents on the basis of rules that have been handed down to them, often through a historical process tracing back to a distant origin.

The collective-acceptance theorist could reply at this point by appealing to the genealogy of institutions: since there is no way to determine that the content of collective intentional states constitutes, rather than describes, a given institution at any given time $t$, the collective-acceptance theorist could set out to show that the institution had its moment of birth when it arose out of collective acceptance, that is, collective intentional states were constitutive of the institution at time $t_0$, and this explains how collective intentional states can be constitutive at any subsequent time $t$. The mechanism, it might be argued, remains the same: collective acceptance exerts its force continuously in determining the nature of the institution. But even if we grant, for the sake of argument, that a given institution did in fact initially arise out of collective acceptance, there would still be another way of accounting for the institution’s nature and development. The content of the institution at any subsequent time $t$, in other words, would be determined and constituted not by collective acceptance at $t$ but by the institution’s genealogical process (with all its intervening modifications). Which is to say that we could appeal to the institution’s history. On this account, it would be in a discontinuous manner that collective acceptance would determine the nature of the institution: collective acceptance would have an essential role in the process of creating an institution at time $t_0$, but for subsequent points in time (at any time $t$) it would be replaced in its constitutive role with a genuinely assertive reference to the

---

original process of creation, to the institution’s overall development, and to the institution’s structure as it emerges from this development.

How should we go about choosing between these two conceptions of institutional genealogy, the continuous and the discontinuous? I believe a comparative assessment favours the latter, for on this conception collective acceptance at any time 
\[ t \]
, different from 
\[ t_0 \], within the group \( a_1, ..., a_k \) and the group \( a_{w}, ..., a_{z} \), could very well be mistaken, because it would describe the content of the institution, rather than constitute it, and this would explain in a quite straightforward manner the assertive nature of institutional discourse. On the continuous view, by contrast, we are asked to take on faith the collective-acceptance theorist’s claim that he or she can judge the real nature of our own beliefs (independently of what we think) and that what we take to be a genuine assertive discourse is only pseudo-assertive, and thus based on an all-embracing collective error.

We have done with the first question, that of collective beliefs as sufficient conditions for the existence of constitutive rules. Let us turn to the second one now, which concerns collective belief as a necessary condition for the existence of constitutive rules. Here I will take a more straightforward approach, by reasoning directly from the previously considered phenomenology of games and asking: Is the existence of a game, conceived as a system of rules, necessarily dependent on there being collective beliefs held by a corresponding community of players? My answer in this case, too, is no—or rather, not exactly, in that there is some truth to this thesis, but it does not capture the whole truth. For, on the one hand, because an institution created through constitutive rules essentially exists as an intersubjective and social phenomenon, constitutive rules must in some way rest on collective acceptance; otherwise, the corresponding institution could not work. But, on the other hand, it is one thing is to say that collective beliefs are necessary for an institution to work and another to say that collective beliefs dictate what the institution is, or that they dictate what the content of its rules is and what the conditions are under which these rules can exist, and that these determinations form an essential part of the institution itself. Suppose that a game is developed but its designers decide not to publish its rules until several years later. Suppose, further, that this publication goes unnoticed for another year or so, until an intrigued player chances on it and circulates the rulebook, and only then does a community of players develop around the game. Here it takes time before a community’s collective acceptance or belief begins to consolidate, and so we ask: Would it be justified to say that the game did not exist before then? If so, how could the intrigued player have recognized and understood the game? How could we say that the soon-to-be community became acquainted with the game (you can only become acquainted with what

---

29 This argument may be challenged as partial because it chooses between the two views without taking all relevant epistemic criteria into account. It might be argued, for example, that a theory of continuity is simpler than, and hence superior to, one of discontinuity. My reply is that, while simplicity may be an epistemic desideratum, it should not be treated as an overriding criterion: yes, the simpler of two theories should be preferred all other things being equal, that is, when the two have the same explanatory power. But when one theory offers a better explanation than another, then simplicity should yield to explanatory power, so long as this does not come at the cost of too much complexity. Although the discontinuous view may be more complex, it fares better than the continuous one in explaining the assertive nature of institutional rules: at least it does not ask us—as the collective-acceptance account of rule realism does—to simply take it on trust that the collective-acceptance theorist has more insight than we do into the real nature of our own beliefs.
already exists)? How could the game developers say, “This game we have here isn’t so bad, after all: Let’s publish it.”?

To be sure, the game as institution could not possibly work in any proper sense until its rules become part of the content of collective intentional states in a community of players, and in this sense collective acceptance does figure as a necessary condition: trivially enough, a game cannot be a game without players, and players must know the rules if they are to be able to play. But this should not be taken to mean that the rules of the game had no content before the community took an interest in the game and began to play: the game clearly had rules, a design, and a point before a community formed around it. We know its rules had content and were constitutive, for otherwise they could not have defined the elements of the game in any sensible way, a way that a human community could understand so as turn the rules into a practice of play: how could the community have understood the game and have collectively formed beliefs about the rules of the game if those rules meant nothing (defined nothing and had no content) and did not exist? And therein, ultimately, lies the paradox of rule realism when it comes to constitutive rules. The paradox, it seems, will not go away until we accept that a community’s collective beliefs cannot be what brings an institution’s constitutive rules into existence, because that existence is actually a precondition: without it, no one could come to know the rules and no community could form playing by those rules (the rules could not define any community of players). It is the institution that makes it possible to engage in the practice of playing, and not the other way around.

The objection here would again be made by pointing out those cases in which institutions do not have a founding moment (with clearly identifiable authors who deliberately got together and designed the institution from scratch) but rather emerge by a gradual process through the community’s collective beliefs, as in the earlier example of a game arising out of the community’s mistaken interpretation of a previously designed game. To this objection I respond with the same objection raised in considering behavioural rule realism. Which is to say that a community’s collective beliefs cannot indefinitely determine the content of the relevant institution’s constitutive rules. At some point, the newly formed game will be found to exist, and from that point on it will always be possible to judge the content of a community’s beliefs—including its collective beliefs—against the game’s original design. This is so because, even if the creative process out of which the game emerged was more indefinite than in the original game, this game (with its original design) nevertheless stands as an essential benchmark against which to assess the new game’s intended content. Hence, what is really necessary for the new game to have content is not that the community have formed beliefs about the game but that the game’s constructive process be considered complete, so that any questions regarding the content of the game will be answered by reference to that process (to the extent that such an answer can be found).

The thrust of my discussion on rule realism should be sufficiently clear at this point. It can be summarized thus: Institutions have a history, and this history is essential in determining their content. Institutions are created by human beings for a reason, and their subsequent development is continually measured in light of these origins. This is because institutions are human artifacts and so have a definite line of development through successive creative stages. It is the whole of this history that fixes the content of an institution, and not simply whatever collective beliefs or behavioural convergence may at any one time for whatever reason take hold where such content is concerned. Institutions, however, are artifacts of a peculiar sort: they are made up of rules, and since rules need to be reasonably clear and fixed in order for anybody to understand and follow them, the same goes for institutional rules and the communities who follow them. To be sure, no institutional practice could exist without this rule-following activity and the accompanying formation of collective beliefs about these rules. But this should not be taken to mean that any such collective
beliefs or any behavioural convergence can entirely determine the corresponding institution’s nature, structure, and content. True, institutions evolve—they change and can be adjusted in light of circumstances, and much of this happens through collective beliefs and convergent behaviour—but what determines the nature and content of an institution is its original purpose in its overall evolution through all its changes. Such nature and content, in other words, lie in the history of the institution rather than in a community’s beliefs and behavioural convergence at any given time.

3. ... But Texts Alone Achieve Nothing

Having done with rule realism, we can now move on to rule positivism, a conception at once easier and more difficult to criticize than the former. It is more difficult to criticize as a general assertion making the simple claim that constitutive rules exist before and independently of the activity they constitute. Rule positivism so framed is in effect nothing more than a derivative thesis in the negative claiming that rule realism is false or, in other words, that constitutive rules do not exist only when accepted by a community as part of the activity they constitute. This broad understanding of rule positivism amounts to an anti-internalist view of social facts and rules, such as that argued by Dan Fitzpatrick in this passage:

Without any independent criteria to establish whether a social fact has occurred or not, it would be impossible to say whether any social or institutional fact has actually occurred. For instance, it would be possible to claim in court that since one had not harbored the relevant we-intention at one’s wedding, that one’s marriage was void despite clear evidence of uttering “I do” at the right moment and signing the register. Or one could claim that one was not actually joining a political party when signing its relevant papers or that one was not actually voting when placing a tick in the relevant box on the ballot paper. But clearly this is not what actually occurs in courts or in connection with other aspects of social reality—not only can we ascertain whether some social or institutional fact actually occurred, we often need to be able to ascertain such facts.30

Rule positivism so interpreted is what one would embrace having rejected rule realism, and in that respect I should consider myself a rule positivist. But when we set out to specify rule positivism as a view unto itself (rather than just a thesis derived by negating rule realism), it becomes easier to spot its flaws.

In what follows, I criticize a specific form of rule positivism, a theoretical model loosely based on Maurizio Ferraris’s “textualism,” for this is to the best of my knowledge the only nonderivative (non-negative) rule-positivist thesis so far developed in social and institutional ontology. The reason I use a theoretical model based on Ferraris’s theory, rather than referring to the theory itself, is that while the theory may be rule-positivist in spirit, there are features of it that make it difficult to so characterize it in a strict sense. Hence the need to give an approximate rendition of it, and in the discussion that follows readers will be able to judge for themselves just how much of an approximation this is.

The main contention of Ferraris’s textualism is that a social object is always created through what he calls a recorded or “inscribed act” (atto iscritto), and hence that the existence of a social object may be considered equivalent to the existence of an inscription (iscrizione). Here is, in particular, the claim Ferraris makes with respect to constitutive rules:

My thesis is that, contrary to Searle’s idea, a social object’s constituting rule is not \( X \) counts as \( Y \) in \( C \) […] but object = inscribed act: social objects are social acts (acts involving at least two persons) characterized by their being inscribed somewhere, in a document, a computer file, or simply in people’s heads. […] social objects, which necessarily have to be constructed, depend on social acts, and it is the inscription of these acts that constitutes such objects.

Now, given that inscriptions (at least in the broad sense understood here of a marking, record, or registration made somewhere) exist independently of what we may believe about them and of how we may choose to act in connection with them, and given that the rules constitutive of a social object are (on this theory) essentially inscribed rules, I have taken this theory to be a form of rule positivism. But several qualifications are needed, because it turns out that Ferraris takes a very broad—and in my opinion overinclusive—view of the term inscription, understood as any “recording [registrazione] having a social value,” and a recording, in its own turn, is understood as a mark or trace (traccia), or any “modification of a surface that can serve as a sign or a reminder.” The crucial point here is that even a neuronal pattern in someone’s brain will qualify as a trace: “Traces occur at the beginning of the process culminating in the assignment of status functions,” this because “the mind is itself made of traces.” On the model of Derrida’s concept of “arch-writing,” Ferraris therefore embraces what appears to be an all-inclusive concept of writing and of texts. Indeed, in his view, arch-writing lies at the very foundation of human thought and activity:


32 Ferraris, Documentalità, p. 90–91; my translation.

33 Ibid., p. 261; my translation.

34 Ibid, p. 250; my translation.

it “includes ritual, memory, animal tracks, and any trace that may be detected by technological means (such as DNA),” as well as gestures.

This is a feature of Ferraris’s theory that I will not carry over into my own model, which thus treats textualism independently of the thesis that even mental records (inscriptions in people’s heads) can be the traces on which basis to determine the existence of a social object. I believe this thesis stands on shaky theoretical grounds and therefore cannot be used for any practical purposes. In fact, despite Ferraris’s several attempts to show that the idea of writing serves as something more than a metaphor in explaining neurophysiological events, and that writing is fully constitutive of thought, I still think this is really a miscast metaphor (however appealing it may be), for at least three reasons. First, traces in our brains are not things we can control; they are not things we do, at least not any sense comparable to our writing on a tablet (performing the act of inscribing words onto a tablet). Traces are artifacts, the products or byproducts of acts, but the same cannot be said of neuronal patterns, which should rather be described as the vehicle or necessary means through which we do anything. Since we could not do anything if we lacked neuronal patterns, we know that such patterns are not something we can “make.” Second, contrary to our ordinary idea of writing—regarded as having some degree of permanence, as something we do “for the record”: *scripta manent*—it is not at all clear that this is the best way to conceive the working of our brain: we could conceive the brain as a pattern storehouse, or we could conceive it as something that reacts to perceptual patterns, as has been argued in research on embodied cognition. On this latter conception, the brain would play an active role (rather than a passive one), and we would accordingly have to rethink the metaphor of writing by framing it as an activity (the act of writing) rather than as something we use or otherwise just store somewhere (writing as an object or medium). Stated otherwise, while we can all appreciate what it means for words and writing (as records or written traces) to be in one way or another fixed onto some kind of support so as to have some “staying power” (*scripta manent*), it is not at all clear what the brain equivalent of that function or feature would be, or even that there is such an equivalent to begin with. And, third, written inscriptions are semantic objects, in that they refer to other things, but neuronal patterns are not: there is nothing they represent, at least not in an ordinary sense of the term. A written trace needs to be recognized in order for it to be a written trace: nothing can exist as a written trace unless someone can recognize it as something that makes reference to something else. No such mental recognition or reference-making need be present in order for something to count as a neuronal pattern, because a neuronal pattern, or rather, its activation, is itself the mental activity we have to undertake in order to recognize something as a trace, as something that needs to be interpreted by bringing it into relation with something else it represents. Neuronal patterns cannot be understood as written traces, because they are precisely the thing we need in order to interpret a written trace—the necessary means through which to interpret the things we interpret—and we do not “interpret” neuronal patterns as objects with semantic content.

Ibid., p. 231; my translation.

Hence, I will rely on a “leaner” account of textualism that does not endorse all the theses of Ferraris’s theory. On this narrower account, institutional objects and constitutive rules do require inscription as a necessary condition (among others) without which they could not exist, but inscription is understood in the ordinary sense of any kind of human-made writing or symbolic marking. There are in particular two sorts of inscriptions that on this narrower account of textualism play a key role: type-establishing and token-establishing inscriptions, the former (iscrizione istitutiva di tipo) being specifically necessary for the existence of constitutive rules, the latter (iscrizione istitutiva di token) for the existence of institutional facts. In the language of contemporary formal ontology, we would say: The existence of a constitutive rule is specifically dependent on a type-establishing inscription; the existence of an institutional fact, on a token-establishing one (where one thing, Y, is specifically dependent on another, X, when it depends on a specific instance of this other thing X rather than on the existence of Xs in general). 38

Now, the question is: Is this narrower form of textualism an accurate description of institutional phenomena? I say it is not. I agree that we cannot determine the content of an institution unless we refer to its creative process, the act through which a social object is constructed. But I do not agree that this creative process must be an act of inscription: it can be, but it does not have to. Indeed, as Ferraris himself concedes, we usually inscribe something in order to record or keep track of it. But an inscription as defined would normally fix (rather than create) an institution: it secures for the institution a degree of fixedness and permanence, and to that extent it plays an essential role in maintaining the institution. But it is one thing to say that an institution in many cases requires an act of inscription (or writing) for its maintenance and another to say that this act of writing (or inscribing) creates this institution (its creation is no more than an act of writing).

It would be unfair to point out the many instances of institutional practices that can take place without inscription, as in the making of promises and other activities based on custom. I say “unfair” because in reality Ferraris explains promise-making, along with many other examples, by means of his previously introduced (and in my opinion overinclusive) concept of arch-writing, and by rejecting this concept, I have in a sense “robbed” his theory of the conceptual means through which to account for nonformalized, custom-like practices. And so I will instead proceed by showing how my narrower account of textualism comes up short even in its attempt to explain highly formalized institutional phenomena, where textualism should find itself more at home.

Consider the example of money. Money—the amount of it in circulation and its constant changing of hands—is something we keep track of by means of inscription, whether on paper (bills, notes, drafts, ledger books, and so on) or digitally (electronic funds transfers, e-money, and the like). But does this mean that the institution of money exists because inscribed in a class of objects? I think not. For this does not explain the general nature and functioning of money as an institution or its specific features as part of a set of practices. There is nothing about inscription that tells us anything about money as distinguished from

other institutions: inscription rather tells us something about our tendency to forget. It is this general and widespread tendency, among other factors, that makes our use of writing almost ubiquitous. We take notes to remember dates, thoughts, and even the feelings we have experienced (we set them down in our journals and diaries). But this does not mean that these inscriptions explain how dates, thoughts, and feelings come to exist. Dates, thoughts, and feelings, and money, too, are different kinds of things that can be recorded via inscription, but taking this to mean that inscription provides the foundation on which basis to explain their existence, or how they come about, does not help us explain in what ways they differ from one another. Inscriptions of very different kinds can trace to a common institution, and, conversely, inscriptions falling within the same class (legal documents, for example) can point to a wide variety of institutions. Hence, if textualism cannot explain what distinguishes institutional objects from other kinds of things we record by inscription, and, all the more so, if it cannot rely on inscription to explain how one institution differs from another, then we have to ask just how textualism can provide us with an institutional ontology.

It should be clear that money is much more than the set of its inscriptions, and the same can be said of games. Consider chess, for example. Clearly, chess has its rules, and these rules are written somewhere. But could we say that if these constitutive rules were unwritten, chess would thereby not exist? I think not. Our making reference to inscription is more an option than a necessity—a very frequently chosen option, to be sure, but still an option. Suppose that chess had only one designer, who conceived the game so as to play it with just three friends. In this tight group of aficionados, the institution of chess could be played using everyday objects as pieces, and in order to know the rules (or their content) it would suffice to turn to the creator of chess (it is quite plausible, in such a small community, that the game’s original design can be amended by a process of successive modifications, and so that the community as a whole could figure as the creator of chess in its successive versions. There is no need here for written rules, or for inscription, or for inscribed objects. But still a circumscribed institution would exist among the group: that of playing chess once a week.

Finally, a textualist account of rule positivism does not explain how institutions require recognition in order to work. Indeed, if we were to exasperate the logic of this conception, we would be forced to say that social reality could survive the total annihilation of humankind—so long as documents (inscriptions) continue to exist. Granted, Ferraris emphatically stresses the point that inscriptions must be used and recognized as such by a human community in order to work, but by insisting on inscription, textualism inevitably sidelines the question of how it is that our recognition of inscriptions can secure the functioning of the relative institutions. In fact, an institution’s dependence on the recognition of its inscriptions is held up as a peculiarity of institutional facts, but the recognition and use of inscriptions is a typical human activity having a role well beyond that of maintaining institutions. So we ask: How does textualism account for the peculiar role of human recognition in institutional ontology? Inscriptions exist in written form regardless of whether or not they are recognized. Hence the further question: What exactly is the difference between an institution existing only in writing and another that is also recognized and practiced? Further, inscriptions can be recognized by readers or users as

See, for example, Ferraris, *Documentalità*, p. 263.
having either a descriptive meaning or a constitutive one, and this shows that the performative nature of constitutive rules does not depend on their being inscribed or on the nature of such inscription: it rather depends, in a sense, on the humans who create and use that inscription.

These points are important because they highlight a risk that all rule-positivist theories incur. As discussed in Section 2, rule realism, for all its faults, still captures a fundamental truth, namely, that although a human community’s recognition of an institution may not necessarily be constitutive of that institution’s existence, it nevertheless stands as an essential condition for its functioning. And so, rule-positivist theories must be able to account for this fact: these theories ground the existence of constitutive rules and institutions in something other than collective beliefs and institutional practice, but they must still account for the essential role of recognition in the functioning of institutions—they have to explain how recognition figures into institutional ontology. I do not see how textualism can do that.

4. Conclusions

In this discussion I first examined the mainstream conception of institutional ontology developed in social ontology: the so-called collective-acceptance view. I argued that this view, if framed as a conception about an institution’s constitutive rules, has several defects on conceptual grounds as well as on phenomenological ones. Specifically, the view cannot explain how institutions can exist to some extent independently of a community’s practices and beliefs, or how institutions can have an objective existence, considering that they very often define the limits of the relevant institutional communities, thus enjoying a degree of independence from those communities.

As an alternative to the mainstream view (collective acceptance), we considered another social-ontological view of institutional ontology, a view that seeks to explain the objectivity of institutions by appealing to texts. But I have argued that this alternative view (a form of textualism) cannot explain how texts need to be recognized if they are give rise to institutional reality, or what kind of interaction between texts and humans we should suppose to hold in the institutional domain.

Hence, the two views here examined suffer from complementary defects. The first fails to explain the degree of objectivity that institutions in fact have in the social context. The second fails to explain in any specific way how the objectivity of institutions can be squared with the fact that institutions require recognition in order to work. The question, therefore, is whether an alternative view can be offered which can solve these problems in a coherent and simple way.