Institutional Externalism

Abstract
Many philosophers regard collective behaviour and attitudes as the ground of the whole of social reality. According to this popular view, society is composed basically of collective intentions and cooperative behaviours; this is so both for informal contexts involving small groups and for complex institutional structures. In this paper I challenge this view, and propose an alternative approach, which I term institutional externalism. I argue that institutions are characterised by the tendency to defer to elements that are external to the content of collective intentions—such as laws, declarations, and contracts. According to institutional externalism, those elements are the grounds of institutional statutes, rights and duties.

I. Does social ontology rest on a mistake?

Social entities are "things" like promises, contracts, presidents, associations, money, debts, corporations, and financial crises. Although there may be social kinds that have a natural basis (candidates are *sex, gender, race, sexual orientation*[1]), most social kinds seem to be essentially dependent on human activities and thoughts. However, the way even the most institutional and conventional aspects of social reality depend on human activities is no trivial matter. Institutional entities are neither flimsy (as individual thoughts and ways of behaving typically are), nor are they always "easier" to modify than physical ones. It may take very little effort to smash a stone into small pieces, but it does not seem to be within the power of any single individual to alter the value of a 5 euro note. A widespread take on social ontology has it that the subjective and yet resilient status of social and institutional reality is given by its dependence not on *individual* mental states but rather on *collective* intentions with shared contents. Even though there is no consensus on how collective intentionality should be analysed, there seems to be substantial agreement on the fact that collective intentions play a central role in explaining why social objects are on the one hand dependent on people, and on the other hand largely independent from individual will and whim[2].

In what follows, I will focus on Searle's version of collective intentionality, and criticise one of its central points. That will allow me to introduce a radically different picture of social reality and the role of intentionality in its constitution. In Searle 1995's well-known account of money, the collective intention of attributing a function to certain pieces of paper—the function of enabling their bearer to exchange them in return for commodities up to a certain established value—is what makes those pieces of paper currency. In this picture, the whole of

---

[1] See Kitcher 1999, and Appiah 2006, Haslanger 2006, Saul 2006, Mallon 2006.
[2] I take the following authors to endorse a version of the view: Gilbert 1989, Bratman 1992, Searle 1995, Tuomela 1995, Hacking 1999, and Tuomela 2002.

1

social reality rests on collective intentions of this sort. In particular, the shared content of the intention determines the nature of the social object to which it gives rise—roughly, what Searle calls the "status function", i.e. the function that an object has in virtue of being collectively attributed a certain status.

I think the explanatory model based on collective intentions with a shared content is flawed[3]. It works (to a certain extent) for simple and informal cases. However, it cannot be applied to the complex variety of social entities that we find in modern societies, in particular to the ones we find in institutional contexts. In the alternative approach that I propose here, contents of shared collective intentions do not determine the fundamental features of institutional entities; collective intentions play a more "neutral" role with respect to the features of institutional reality. Roughly, institutional contexts are characterised by a shared intention to defer to elements of the context with regard to what determines the nature of institutional reality. Such elements are laws, contracts, and the like (what I will call "enactments"), which are in principle publicly accessible, and are regarded as valid and binding because they are produced according to an established procedure; thus, *not* in virtue of their content. In what follows, I start with (§II) highlighting Searle's implicit assumption of continuity (AC), which I take to embody a widespread stance in contemporary social ontology. I then present the principles of my alternative view (§III). In the following two sections (§IV and §V), I discuss two problems that follow from the continuity assumption, and then (§VI and §VII) I argue that Institutional Externalism resolves them and has further merits. In §VIII I draw conclusions.

II. The Assumption of Continuity

John Searle's theory of status function is one of the most developed forms of social ontology on the market, and one that shares many widely accepted assumptions with other theories. According to Searle, social entities depend on people in the sense that they have properties that they would not have if people did not exist. To pick up on the example of money: What makes it the case that a piece of paper can function as a 5 euro note? It is certainly not its microphysical structure or some other intrinsic property. It is the fact that there is a community of people that recognise it as having precisely that function (although there may be physical constraints on the social function: a 12-ton coin would be very impractical). Ontologically speaking, then, social objects depend on people for both their existence and their features. If there were no people, there would be no society, and if there were not certain behaviours and thoughts (performed or entertained by people) a society with such and such features would not exist. According to this theory, we find this sort of dependence all the way down and all the way up in social reality:

"Collective mental phenomena of the sort we get in organized societies are themselves dependent on and derived from the mental phenomena of individuals. This same pattern of dependence continues higher as we see that social institutions such as governments and corporations are dependent on and derived from the mental phenomena and behaviour of individual human beings"

---

[3] Thomasson 2002 and Khalidi 2013 have raised doubts about the generality of Searle's idea that social kinds require collective attitudes towards them in order to exist, but for different reasons from the ones that I discuss in this paper. In particular, they do not question the validity of the model in institutional contexts. Guala 2010 takes the model to be a (useful) idealisation. See notes 8 and 9 below.

(Searle 2010: 4)

In a sense, this continuity of dependence on individual intentions from simple collective behaviour to complex institutions is obvious. What would be the alternative? Maybe some caricature of the "Hegelian spirit" which is embodied in the higher tiers of social reality and whose objectivity is utterly independent from what the individuals composing the society think and do? It seems more in line with a naturalist methodology (even a very moderate or "liberal" one) to accept that there is no substantial difference between the informal collaboration among people in small groups and the kind found in much more complex situations, such as those involving governments issuing laws, contracts, and institutions. However, this trivial reading of Searle's passage is not what is meant by the idea that collective intentionality is the ground of social reality.

Searle's approach requires not only (a) that all intentionality should be in individual minds or brains, and thus that also collective intentionality is at bottom a psychological phenomenon that concerns the individual (independently from the issue of whether its form is irreducibly plural or not), but also (b) that collective intentionality is "the fundamental building block of all human social ontology" (Searle 2010: 43, italics mine). That is to say, it is in virtue of collective intentions and their content that there is a domain of social objects that possess characteristic features and have, in some sense, objective existence. I will label these two basic requirements the thesis of individualism (TI) and the assumption of continuity (AC).

(TI) All intentionality is in individual brains

(AC) Social phenomena are determined by collective intentions and their shared contents in small groups as much as (and in the same ways as) in complex social situations

Many philosophers agree with the gist of Searle's position, at least with respect to (AC)[4]. Others, like Ruben 1989, Thomasson (2002, 2003, and 2009), Khalidi 2013 and Guala 2012 disagree on whether the intentions involved in the construction of a social kind K must always involve the concept of K, and argue that social reality may show much more "opacity" than Searle's account entails. Here, I intend to criticise the grounding role of collective intentionality for institutional reality, which is elicited by (AC). I will remain neutral with respect to (TI). I start by presenting my alternative proposal, which I term Institutional Externalism. I then discuss the flaws of (AC) and argue that Institutional Externalism can overcome them.

III. Institutional Externalism

I take (AC) to fail because in institutional contexts the role of collective intentions is *not* that of determining the nature of institutional phenomena. Insofar as institutional contexts are part of complex social situation, thus, (AC) fails.

---

[4] While (TI) is explicitly maintained by Searle (see Searle 1990: 406), (AC) is only implicit in his writings (but passages analogous to the one quoted two paragraphs above are not difficult to find in Searle). Gilbert 1989 and 2006 reject (TI).

Examples of social facts that require institutional contexts are the fact that a certain person is the chief executive officer of a certain company, that a certain action is a tax crime, that a certain corporation exists, or that I have acquired certain duties by signing a contract. In other terms, facts about the existence of chief executive officers, tax crimes, corporations, marriages, and the like, are not grounded on the shared content of collective intentions. The nature of institutional entities depends rather on *enactments* that are collectively regarded as valid. I will use the term "enactment" to cover things such as laws, declarations, and contracts, whose contents are usually written in documents, and are always external to the content of collective intention, in the sense of being elements of the context that are in principle publicly accessible. Institutional Externalism can be summed up in two theses:

(Institutional Realism) Facts about the existence of institutional entities are partly grounded on facts about material individuals and events and partly grounded on facts concerning enactments in institutional contexts[5].

(Content Deference) Institutional contexts are characterised by the shared intention to regard the enactments that are produced according to established procedures as valid, namely to defer to such enactments in order to establish the socially binding contents[6].

Consider two examples: the case of a tax crime, and that of a marriage. Suppose that an individual S carries out at a time t a certain financial operation which in an institutional context C is a tax crime. According to Institutional Realism, the fact that there is a tax crime (committed by S at t) is grounded on facts about what S did at t (the financial operation that she carried out), and facts about the tax law (an enactment) in context C. According to Content Deference, in C the enactment which establishes that S has committed a tax crime (i.e. a certain tax law) has been produced according to an established procedure that is collectively regarded as producing valid enactments. That is to say, the content of the enactment (which says that whoever carries out certain financial operations commits a tax crime) is collectively regarded as binding *not* because it is the content of a shared intention, but because people in the institutional context defer to what enactments produced according to established procedures declare to be legally binding. The case of marriage is slightly more complex, because it (usually) requires two enactments: a marital law and a marriage certificate. Suppose individuals S and S' do what it takes to get married in context C: let's say, undergoing a marriage ceremony and signing a marriage certificate. According to Institutional Realism, the fact that they are married is grounded on facts about what they did in the ceremony, and facts about both the marital law and the marriage certificate they signed in context C (*viz.* facts about two enactments). According to Content Deference, in C the marital law has been produced according to an established procedure that is collectively regarded as producing valid enactments. The binding content requires there to be a further enactment (the marriage certificate), which is produced according to an established procedure. The content of the certificate, usually in the terms specified by the law, is thus also collectively regarded as binding.

---

[5] On grounding and non-causal determination, see Correia and Schnieder 2013, Sider 2011, Fine 2001. Certain "abstract" institutional entities — such as corporations, debts and electronic money —require only enactments as grounds. See Thomasson (2002: 282-3; 2003: 589; 2009: 548) and the discussion on "free standing Y-terms" in the exchange Smith and Searle 2003.
[6] The idea of deference here is taken from Putnam 1975 and Burge 1977.

Therefore, an enactment can depend on others, and the detail of how the content that is regarded as binding is specified may involve several different enactments (often contracts and declarations refers to laws instituting the possibility of issuing them). In the following paragraphs, I will spell out Content Deference in more detail, and argue that it is the correct way to understand the role of collective intentionality in institutional contexts[7].

IV. Against the Assumption of Continuity: the Explanatory Gap

In his 2010 book, Searle claims that collective intentionality manifests itself in two ways. The first, cooperative behaviour, involves action; the second, collective recognition, involves belief. Cooperative behaviour does not require that a complete intentional content be shared: the only part of the intentional content that has to be shared in order for cooperation to occur is the goal. If we cooperate to play a duet, I may play the violin while you play the piano: our intentional contents will be different, but we will share the same goal (viz. playing a duet). Collective recognition requires a more complete shared content: an attribution of a particular function to something or someone. We have collective recognition when people in a community recognise that something has a certain function, as, for instance, when everybody agrees that pieces of paper produced in certain ways have a given monetary value. The two forms of collective intentionality can work in tandem. Consider a simple case. Four friends decide to go camping for the weekend. They cooperate to assign responsibilities for setting up the tent (two people), gathering wood for the fire (one person), and cooking (one person). As a result of this cooperative assignment, there is collective recognition of the role that each person has. According to (AC), much more complex situations than this one work basically in the same way, namely through cooperation to accomplish shared goals, and collective recognition of status.

Although cooperation is the fully fledged form of collective intentionality, according to Searle, collective recognition is somewhat more fundamental to social reality, in that it yields the minimal requirement for there to be a social domain. And, given (AC), it should come as no surprise that collective recognition is also the minimal condition for there to be institutions, which are the most complex forms of social organisations:

"[...]the existence of an institution does not require cooperation but simply collective acceptance or recognition" (Searle 2010: 58)

Let us then focus on collective recognition and institutions. Given (AC), what is claimed about institutional reality is just an extension of what happens at less complex levels of social organisation. We say that there is an institution X in a given society when we find institutional facts, to the effect that certain people (and/or objects) have properties that they would lack if there were no X. The thesis is that institutional facts are grounded in collective beliefs: in every case in which we have an institution, people and objects possess the associated social

---

[7] Although, to my knowledge, institutional externalism is an original position in philosophy, I think the "spirit" if not the "letter" of it is often to be found in the social sciences. In economics (e.g., Hayek 1943, and Schotter 1981), but also in the study of social cognition, in which we often find the idea that one key role of institution is allowing for computational economy by providing external support Clark 1997: Ch. 9, Aoki 2011.

properties in virtue of a collective recognition of such properties.

I have two main objections to this thesis. Firstly, taken literally, it leads to implausible consequences regarding the shared content of such grounding beliefs. Assuming that in the toy-example of the four friends on the camping trip it works, it is highly implausible to maintain that most people (in our society) hold *detailed beliefs* about the social role or function of most of the objects and persons surrounding them. (If you doubt this, ask people what are the duties and powers of the chief executive officer of a company, or what they have agreed to by clicking on the last update of their music software). Still, if the social function is grounded on a collective attribution, then only a feature of the shared content can account for a feature of the social entity. If we acknowledge that the social entity has a property that does not result from the collective attribution (and is not just grounded on its physical constitution, if there is one), then we need to revise the standard account somehow, because its generalisation from simple cases fails. The challenge that any approach based on (AC) has to face, then, is to fill the explanatory gap between the rough and imprecise beliefs (if in fact there are any at all) that most people have with respect to most institutional entities around them, and the very fine-grained system of roles, statues, duties and rights that we find in institutional contexts.

One may have the feeling that this critique is overhasty. It is true that we cannot extend the standard account from simple cases to more complex ones without some adjustments. However, how these adjustments should be made is quite straightforward. The collective intentional contents on which social reality depends according to the standard account are not necessarily wholly explicit. Even though it is often convenient to describe these intentional contents in terms of explicit propositional contents, as attributions of a certain function to a certain entity, their psychological status is not necessarily that of a propositional belief. It is feasible to view them as a set of dispositions and capacities (cf. Searle 1995's notion of "background capacities"). In particular, in institutional contexts, I do not need to have detailed knowledge of the functions that I attribute implicitly to the chief executive officer of a company in respect to his status as CEO, in so far as other people have it and I have a *disposition to defer to them* for such knowledge. We can call the people with the relevant knowledge concerning a given status function "the experts"[8].

This revised view will not suffice. It is implausible to maintain that each institutional status requires *specific experts* (or people otherwise worthy of being deferred to). Maybe for the category of permanent residents the appointed officers

---

[8] See Tuomela 2002: 200 and Guala 2010: §7. Khalidi 2013, too, seems to believe something along these lines, when he argues that institutional kinds (such as *permanent resident* and *prime minister*) are the ones that fit Searle's model best. "Searle's thesis holds more nearly when it comes to social kinds of a purely conventional nature, that is kinds whose associated properties or conditions of membership are more strictly laid out in a set of rules or laws. [...] To be sure, it is not the attitudes of members of society at large that determine one's status as a permanent resident, for this is usually determined by the attitudes of *officials of the state*, who are informed by the appropriate laws and statues" (Khalidi 2013: 7; italics mine). While I take most of what Khalidi 2013 says to be correct, I think he is misconstruing the role of the "appropriate laws and statues" as a mere prompt for the attitudes of the officials. For reasons that I give in §V-§VII, I take laws and contracts (and enactments in general) to have a much more substantive role. I also suspect that underlying Khalidi's tenet that institutional reality fits the model there is something like (AC) at work. Note also that Searle himself describes the roles of written documents as merely epistemic (1995). Ferraris 2012 discusses at lenght the crucial role of documents and writing for social reality. See also Ferraris and Torrengo 2013.

count as the experts, and we defer to their attitudes to establish who counts as a permanent resident in a certain area; but what about all the cavils involved in my status as contractor that I obtain by clicking at the bottom of a long agreement on a website, or those involved in my job contract? Who are, in those cases, the experts who possess the relevant attitudes towards me (or even a more generic attitude towards whoever signs or clicks)? It also seems implausible to maintain that each institutional status requires *specific deferential beliefs*. To *whom* am I deferring when I sign a contract? It might be claimed that the deferential beliefs that ground institutional kinds are very general (indeed, they are non-specific dispositions to defer) so it is plausible to maintain that nearly everyone has them. Yet this would not be an answer, but merely a way of rephrasing the explanatory problem facing any view according to which institutional entities are grounded on the content of shared intentions: if the shared content is *general*, how can it determine the *specific* variety of institutional reality? A shared intention with a general content of defer may be a necessary conditions for the existence of institutional entities, I agree, but that simply means that the account according to which institutional entities are determined by the shared content of collective intentions require a (substantive) integration. Institutional statuses are complex sets of functions, constraints, rights, and obligations that cannot be real if their only purported ground comprises general beliefs or dispositions. Elaborating an account of how the contents to which we defer play a role in grounding social reality seems a promising way to overcome this predicament. However, as will soon become clear, by so doing we will discover that the fundamental assumption of continuity has to be drastically reconsidered, and in a way that leads us to institutional externalism.

V. Against the Assumption of Continuity: Collective Mistakes

The second objection to (AC) is that it does not seem to leave room for mistakes about social statuses and about social facts in general. Even mistakes on the part of individuals are problematic for a naïve reading of the thesis. Consider the following example: I promise to take you to the cinema tomorrow. Tomorrow, though, you think that the day before I promised to take you to a high-class French restaurant. How are we to straighten things out? We seem to have the intuition that I am right (I did promise to take you to the cinema and I did not promise to take you to the restaurant) and you are wrong (you have misremembered the promise or you are making some other sort of epistemic mistake). However, Searle's account predicts a stalemate here. The status of my obligation is unclear, because there is no shared collective intentional content (assuming you are not just trying to tease me), and so nobody is right.

One could counter-argue that in so far as the existence of rights and obligation is concerned, a stalemate is what we should expect. If you and I are the only ones who have partaken in the act of promising, then as soon as you (or I) forget it or fail to recognise it, the promise is no longer binding and so no longer exists. That may be true for informal contexts involving only a few people. The situation is different if many people have heard me promising, and even more different if many people are ready to see to it that we keep the promise. In that case, a plausible reading of (AC) predicts that you would be mistaken because the majority (or the relevant "experts") think that you are mistaken; exactly as, in

general, you have a certain social status only because the majority (or the relevant "experts") thinks that you have it.

I agree that for non-institutional, informal, social situations that is correct, and that in such contexts individual mistakes are compatible with (AC). But what about institutional contexts? Consider the case of a collective mistake, for instance a football match of the national league between team A and team B. Team A wins the match 3-1, but for some strange reason everybody comes to believe that team B won (perhaps everyone in the stadium suffered a drug-induced flashback to the previous year's game, when team B won 3-1, and then the incorrect result was broadcast nationwide over the television and radio). Would not everybody be mistaken? Yet how can this be the case, if the social status of the match depends only on collective attribution and everybody believes that B won? If (AC) holds, they are not mistaken; team B has won. Yet this is wrong. If everybody came to know that A scored three goals while B scored only one, but they believed the reverse happened because their beliefs were drug-induced, everybody would realise that they were wrong, and not that the result had changed! In an institutional context, the final results would be *corrected* after that the mistake has been discovered.

One might defend (AC) by claiming that in the case of a football match, the status of the final result is not entirely a matter of what people believe. It also depends on how many goals have been scored by each team (even though, of course, what counts as a goal is, in turn, a matter of what people believe). Thus, people may have false beliefs concerning the historical facts and thereby have collective false beliefs about who won. It is quite plausible to think that there is a sort of implicit deference to the relevant historical facts here, in the sense that what the historical facts are partly determines the nature of the institutional facts. However, collective mistakes may involve basically every kind of social status, at least in so far as we are in an institutional context in which we can collectively decide to defer to what is written on an external accessible support to establish what is socially binding. For instance, by signing a contract I (usually) accept a large number of duties and assume a large number of rights regarding which I, just like everybody else, may entertain false beliefs. Yet it is not my beliefs that establish what my duties and rights are, nor the beliefs of the collective I am part of, but the content of the written text of the contract. I defer my knowledge to it, because it contains all the information that determines the status function relevant in the situation. And I defer my knowledge, because I defer to its content as to what determines institutional reality. In general, it should always be possible for there to be a discrepancy between what someone (or everyone) believes about the institutional status of an entity x, and its actual institutional status. Yet an account of social reality based on (AC) has difficulty providing an explanation of such a possibility, unless it is augmented with the idea that in institutional contexts we defer to *external* contents. This latter consideration takes us beyond any account

based on (AC) towards Institutional Externalism[9].

VI. Intentions and Established Procedures

According to Searle's widely influential model of social ontology, the collective recognitions on which social reality is grounded are given by cooperation among individuals. Collective recognitions require a shared intentional content: the assignment of a status function to something or someone. As I have argued, the complexity of many status functions in our modern societies and the possibility of collective mistakes about the social status of something cast doubts on the tenability of such a model. For simple situations of collaborative behaviour among members of small groups the model may work, but to account for more complex situations it needs to be revised. I think that the problem here is deep, and that the revisions that need to be made go far beyond small adjustments that would make (AC) still tenable.

The intentions that are required for there to be complex social entities such as institutions do not have the form of the collective recognitions which they possess in simpler cases. Consider again the simple example of a promise: I promise you that I will pay you 5 euros by tomorrow. By this act, which requires a community of at least two persons, a social entity is brought into existence: a promise. You and I are affected by the promise in certain ways: I have an obligation to pay you, and you have the right to request 5 euros from me. In other words, the promise has a binding power. Now, what is it that makes the promise binding? The fact that the parties jointly agree to keep the promise is a precondition for the existence of the social constraints that arise from the act of promising. However, facts regarding who is bound, and how and with respect to what they are bound, depend on the content of the social act that the two parties have performed. The content tells us what has been promised to whom, and thus determines the nature of the bond between the parties.

In simple situations the relevant content usually coincides with that of the intentions of the parties involved in the promise (or their community, or a relevant part thereof), and those intentions have the form of a recognition of a status function. Yet, even in the case of a promise between you and me, if we write

---

[9] It should be clear by now that I am not discussing here the fact that there are social entities that do not require collective recognition *of their kind* in order to exist. Amie Thomasson argues that the cases of racism, gender discrimination, and superstition, but also those of economic crises, recessions, and artefacts are like this. "A society with widespread collective beliefs that supernatural occurrences can be brought about by performing certain rituals, automatically is a society that has the additional social feature of being superstitious, without the need for anyone to accept any constitutive rules about what 'counts as' superstitious. In each case, the mere existence of the relevant collective intentions, perhaps in the relevant context, is logically sufficient to ensure that the relevant social fact of kind S obtains, without the need for anyone to have any thoughts about facts of kind S at all." (Thomasson 2002: 287-8). I do think that such an account of the epistemic "opacity" of many aspects of social reality is on the right track. However, Thomasson (2003: 588-92) also thinks that there cannot be massive mistakes or ignorance with respect to social kinds, not even with respect to highly complex institutional kinds. Again, I suspect that something like (AC) is at work in this case as well. Note that Thomasson is not denying that we cannot make interesting discoveries about social kinds (for instance, concerning their unintended consequences). Guala 2010 argues that we can make interesting empirical discoveries about the causal properties of institutional kinds; and given that the nature of institutional kinds is determined by their causal properties, and not necessarily by their collective acceptance, we may be massively mistaken about them. Again, I am utterly sympathetic towards the idea that there are many elements in social reality that does not require *any* form of collective recognition, and most of them involve institutional elements, at least in contemporary societies.

down that I promised to pay you 5 euros by tomorrow and we agree that the content of this inscription is binding, even if we both forget what has been promised, there will be a way, in principle, to settle the issue. Therefore, if we both agree to defer to an external content, the persistence of the social bond and the nature of the social constraints at issue will be determined by this external content. According to Institutional Externalism, this is what usually happens in more complex situations. In institutional contexts, the content that determines all the social constraints and socially relevant facts is not that of the actual intentions of the participants of the social act which gives rise to the social object. As we have seen, people often do not have detailed beliefs about many status functions, and they can be mistaken about them. Institutional situations require deference to publicly accessible contents, and not a specific collectively shared content. Note that here I am *not* trying to substitute a kind of continuity based on shared intentions with the idea that external supports are essentially involved in any social situation, no matter how informal[10]. Suppose that I make the promise o pay you 5 euros by tomorrow, and we both agree to write it down. Suppose I do the inscription with a computer and it ends up saying I pay you 500 euros (while we thought it says 5,00 euros). Do I now owe you 500 euros as that is what the inscription says? If the context of the promise is non-institutional, it does not seem so, even if we said that the content of this inscription is binding. When the external content is not essential to the existence and nature of the social entity, we almost never give up our interpretative power to documents. However, in a context where the writing of the promise is part of an institutionalized procedure, then — once (and if) we realise the mistake — we *must* undertake some further action to change the external content in order to change the institutional facts. It is not that our interpretive power or intentions are irrelevant, but *they are not what constitute the features of the institutional reality that we inhabit*. In other words, I do take discontinuity between informal contexts and institutional contexts very seriously. The difference between the two kinds is not just a difference in the way social reality comes to be, it is a difference in the kind of ground that it has.

The difference between a view based on the shared content of collective intentions and Institutional Externalism is profound. On the first view, what determines the nature of social entities *tout court* — hence also in the case of the institutional sphere — is a shared intentional content. The elaborate social statuses that we find in complex social structures are grounded in shared beliefs about such social statuses. As I argued, there is an explanatory gap here between the rough grounding of shared intentions and the actual fine-grained complexity of the grounded institutional phenomena: generally, people have only rough and general beliefs about complex status functions. In contrast, on the proposed view, in institutional contexts the intention that is required for the existence of a status function F is not necessarily a collective recognition of the status function F—not even by a designated group of "experts". Rather, the shared intentional content consists in deferring to the content that has been produced by an established procedure. Thus, social intentions of complex structures have to be general,

---

[10] Such a position would have a problem in explaining how enactments have a linguistic content. In so far language has a institutional status of some sort (which seems difficult to doubt), any linguistic content would require a *further* external support that grounds it. In such a scenario, regress (or circularity) is unavoidable. But, again, I maintain that there is *discontinuity* between informal social contexts and institutional contexts, precisely with respect to what grounds social reality in each. While informal contexts may not require external factors as sources of binding contents, but may entirely rely on shared intentional contents, institutional contexts essentially involve deference to established procedure and external contents.

because we need a nonspecific intention to regard the production of a certain publicly accessible content as valid and binding, so that the status function will be determined by the specific content produced, regardless of whether or not it is entertained as a shared content by the parties involved (or any other person in the social context, for that matter).

On the proposed view, the collective recognition that grounds institutional reality is a form of cooperation, but in a much weaker sense than Searle proposes. On his view, the social world, composed of social entities, exists only if there are collective intentions with a shared substantive content, such as that a piece of paper possessing certain features is a 5 euro note. Even if, strictly speaking, a common goal is not required for collective recognition, the fact that the shared content has to do with the specific features of the social entity makes collective recognition and cooperation two sides of the same coin, as it were. In a sense, a shared assignment of status function is a limiting case of a common goal (Searle 2010: 65-8). On the proposed view, the existence of institutions requires only that the individuals in a collective share the general belief that the enactments produced according to certain established procedures determine social constraints, rights, duties, and the like. The shared part of the intentional content is general, not only in the sense that it does not contain any specific attribution of social functions, but also in that it does not concern any specific procedure. In order for there to be institutional kinds and entities, it is required only that we defer to enactments produced by the procedures that are established in the context in which we find ourselves. How the content of this shared intention may be described in detail is probably at least partly an empirical issue, which is connected to the wider problem of social coordination. According to the general working hypothesis, which procedures are regarded as producing valid enactments is largely a contextual matter, and it may depend on the cultural background. What the thesis predicts is that the relevant shared intention is essentially deferential in nature, and that its target is the content produced by a contextually established procedure.

VII. Institutions and Non-collaborative Behaviour

Collaborative behaviour of a more substantive kind than collective recognition of the validity of certain procedures to produce documents or other forms of binding contents is a widespread phenomenon in society, and it is very important and valuable in many social contexts. Searle's view takes cooperation (especially in its degenerative form of the collective recognition of status function) as the basic building block of social reality. However, as I have argued, while this view may work for simple cases, it is wrong to generalise it to complex cases. A further objection to any view based on (AC) is that collective recognition of status functions seems to require acceptance of the assignment of status functions. Discussion of this point will allow me to clarify why deference is an essential aspect of institutional reality.

According to Searle, there is an association between recognition and acceptance: the recognition is shared by being accepted by at least some individuals of the relevant group. However, there certainly can be cases in which we are involved in social situations of which we do not accept. A defender of (AC) can reply that

shared recognition may come with apathy, distress, or even hatred towards the attributed status at least in so far as there is some higher-order goal that is shared (e.g. the cohesion of society)[11]. If I am playing chess with you and you are winning, I will recognise that you are winning even if you are humiliating me, in so far as I am willing to play with you (i.e. I aim to win, but I also have the higher-order goal of playing with you, and the latter is what we share). Or, another example, if I am willing to live in a democratic society, I will recognise the social status and power of the elected president even if I strongly disagree with their political views.

Now, is it really plausible to suppose that a specific higher-order collaborative intention is always at work in every case in which people behave non-collaboratively towards the institutions? Take a prima facie simple case, such as getting fined for parking. Suppose that I think that the officer has made a mistake in fining me. I can refuse to consider this particular fine as justified, while continuing to regard the practice of fining (and other "useful" forms of social punishment) as generally justified. However, lack of justification with respect to a specific fine is *not* tantamount to believing that the fine lacks validity. I may not at all be happy to recognise its validity, but this is what I am doing if I protest—it would be pointless to protest if I thought it to be invalid.

In the proposed view, the general intention on whose ground non-collaborative social behaviour can occur does not need to be a higher-order one. Generally speaking, complex situations require only the first-order intention to consider validating the established procedures we find around us. Consider again the case of the fine. If I think that the officer has made a mistake, I will try my best to have the fine dismissed; that is, I protest the validity of a certain document, the official piece of paper that the officer has issued to me upon which the fine and the circumstances that warranted it are recorded. Such a protest can only be upheld by appealing to another validating procedure, which typically will require me to provide evidence for my claims. Indeed, even if I consider the whole practice of fining to be unjust, I do not consider it as lacking validity if I engage in some form of "protest" against it. Again, it would be pointless to protest if I do.

Yet one may have the feeling that the proposed view is too minimalist. People do not always behave so apathetically towards institutions and society. Even granting that the content of collective recognition is often too simplistic and general to ground sophisticated institutional functions, and that we need publicly accessible records to flesh out the details of status functions, the content of the shared intentions that we find in society is relevant for the existence of the corresponding social object or function. If nobody cared about a certain social function, it is difficult to see how this function could affect our social life, even when there is a document which describes it in full detail and has been issued by an established procedure[12].

This objection is barking up the wrong tree. The thesis that institutional functions are grounded on enactments produced by established procedures to which we

---

[11] Searl 2010 addresses explicitly this concern and he seems to suggest a solution of this kind. He explicitly says that recognition does not require *approval*, but if sharing a higher order goal suffices for acceptance, then it does require *acceptance*.
[12] Khalidi 2013 makes a similar point in discussing the case of a "war" that is officially declared but never actually fought.

defer is compatible with the fact that we do care about what these functions are, and indeed also about the established procedures as such. The reason we are interested in them is that they can, and often do, have important consequences for us. To be sure, in modern societies, the established procedure of producing laws in parliament affects the lives of everyone. The picture is also compatible with collectively shared beliefs having a crucial causal role in explaining social changes over time. For instance, if nearly everybody disagreed with a certain content produced in an established procedure, which has far-reaching consequences (e.g. a law), it is likely that people will start to act in order to change this content (either inside or outside the boundary of what is institutionally permitted). Or, more dramatically, if at a certain point the majority of the population start to share a hostile belief against an established procedure (like the parliamentary procedure of issuing laws of a given state), it is likely that they will collectively aim to disestablish the institution in question. Besides, procedures can be established through their coercive imposition, and small minorities can affect directly established institutions (as a consequence, for instance, of terrorist actions).

Hence, I am not denying that shared beliefs with a substantive content (as to what the social and institutional feature of a given situation are or should be) or substantive collaboration with respect to a common goal are often crucial aspects of social reality. However, common goals and shared beliefs emerge more decisively in the account of the (broadly speaking) causal mechanisms that underlie many aspects of society, rather than in the account of the peculiar *ontological status* of the institutional sphere. The existence of an institutional sphere requires only the presence of established procedures; it does not require the sharing of more substantial goals or beliefs. Besides, within an institutional framework, we find both cases of harmonious cooperation between people sharing a common goal, and cases of difficult negotiation over conflicting interests. Modern societies allow for both. Sometimes there is substantive collaboration towards achieving a common goal, such as when we want to play a duet and we decide that I will be playing the violin and you the piano. However, at other times, there is agreement only on the very "general rules of the game". Substantive collaboration can be a valuable thing in certain circumstances precisely because it is not the basic form of social behaviour to be found in every circumstance. A duet is a wonderful thing, just as society in many cases is not.

VIII. Conclusions
In this paper I argued that most of contemporary social ontology rests on the mistake of providing an "internalist" account of institutional reality. The thesis that the content of shared intentions determines social functions, roles, and the like has some plausibility with respect to informal and non-complex social contexts, but faces two serious problems when it comes to more institutionalised aspects of society. Firstly, it has great difficulty explaining how often very imprecise and general shared content can ground complex and often detailed institutional profiles; and secondly, it cannot account for the possibility of collective mistakes about social facts. I propose an alternative view, which I term institutional externalism, according to which the only relevant shared intention in an institutional context for the existence of institutional entities is that of deferring to established procedure as to what contents are socially binding, namely as to what grounds facts about institutional reality. The view does not suffer from those

problems and it also provides a framework in which the role of substantive collaboration and substantive shared intentions in society can be correctly assessed.

Acknowledgments

References
Appiah A. K. (2006) "How to Decide if Races Exist", *Proceedings of the Aristotelian* Society, CVI: 365-382

Aoki, M. (2011) "Institutions as Cognitive Media between Strategic Interactions and Individual Beliefs", *Journal of Economic Behavior and Organization* 79: 20-34.

Bratman M. (1992), "Shared Cooperative Activity", *Philosophical Review* 101, 2: 327-41

Burge T. (1977) "Belief *De Re*", *The Journal of Philosophy*, 74:338-62 Philosophical Papers. Oxford, OUP: 261-320.

Clark A. (1997) *Being There: Putting Brain, Body, and World Together Again*, Cambridge, Mass., MIT

Correia F. and Schnieder B. (2013) eds., *Grounding and Explanation*, Cambridge University Press

Ferraris M. (2012) *Documentality: Why It Is Necessary to Leave Traces*, New York, Fordham University Press (Original Edition in Italian 2009).

Ferraris M. and Torrengo G. (2014) "Documentality: A Theory of Social Reality", *Rivista di Estetica* 57: 11-28

Fine K. (2001) "The Question of Realism", Philosophers' Imprint, 1: 1–30.

Gilbert M. (1989) *On Social Facts*, London, Routledge

Gilbert M. (2006) *A Theory of Political Obligation: Membership, Commitment, and the Bonds of Society*, Oxford, OUP

Guala F. (2010) "Infallibilism and Human Kinds", *Philosophy of the Social Sciences* 40 (2010), pp. 244-64.

Hacking I. (1999) The Social Construction of What?, Harvard University Press.

Haslanger S. (2006) "Philosophical Analysis and Social Kinds: What Good Are Our Intuitions?", *Proceedings of the Aristotelian Society* 80 (Sup. Vol): 89-118.

Hayek, F.A. (1943) "The Facts of the Social Sciences", *Ethics* 54: 1-13.

Hudin J. (2010) "Can Status Functions be discovered?", The Journal for the Theory of Social Behaviour.

Khalidi M. A. (2013) "Three Kinds of Social Kinds", *Philosophy and Phenomenological Research*

Kitcher P. (1999) "Race, Ethnicity, Biology, Culture", in L. Harris (ed.) Racism, Humanity Books, pp. 87-117.

Mallon R. (2007) "A Field Guide to Social Construction", Philosophy Compass 2(1): 93-108.

Putnam H. (1975) "The meaning of "meaning"", Mind, Language and Reality, Cambridge, Cambridge University Press: 215-271

Ruben D. (1989) "Realism in the social sciences" in H. Lawson and L. Appignanesi (eds.) *Dismantling Truth*, London, Weidenfeld and Nicolson: 58-75

Saul J. (2006) "Philosophical Analysis and Social Kinds: What Good Are Our

Intuitions?", *Proceedings of the Aristotelian Society* 80 (Sup. Vol): 119-144

Searle J. (1995) *The Construction of Social Reality*, New York, Free Press.

Searle J. (2010) *Making the Social World,* Oxford, Oxford University Press.

Smith B. & Searle J. (2003) "The Construction of Social Reality – An Exchange", *American Journal of Economics and Sociology*, Vol. 62, No. 1: 283 – 309.

Sider T. (2011), *Writing the Book of the World*, Oxford, OUP

Searle J. (1990) "Collective Intentions and Actions", in Intentions in Communication, P. Cohen, J. Morgan, and M.E. Pollack, eds. Cambridge, Mass. Bradford Books, MIT Press, 1990.

Smith B. & Searle J. (2003) "The Construction of Social Reality – An Exchange", American Journal of Economics and Sociology, Vol. 62, No. 1: 283 – 309.

Schotter, A. (1981) *The Economic Theory of Social Institutions*, Cambridge, Cambridge University Press.

Thomasson A. (2002) "Foundations for a Social Ontology", *Protosociology: An International Journal of Interdisciplinare Research*, 18-19: 269-290

Thomasson, A. (2003) "Realism and Human Kinds", *Philosophy and Phenomenological Research* 68: 580-609.

Thomasson, A. (2009) "Social Entities" in R. Le Poidevin et al (eds.) *Routledge Companion to Metaphysics*, London, Routledge

Toumela R. (1995) The Importance of Us: A Philosophical Study of Basic Social Notions, Stanford Series in Philosophy, Stanford University Press

Tuomela R. (2002) *The Philosophy of Social Practice: A Collective Acceptance View*, Cambridge, CUP

Author Biography

Giuliano Torrengo is an assistant professor at the University of Milan. He is the founder and coordintator of the Centre for Philosophy of Time, and his main interests are in analytic metaphysics, philosophy of language, and social ontology.