

Ylikoski, Petri & Pekka Mäkelä 2002: 'We-attitudes and Social Institutions', in *Social Facts and Collective Intentionality* (ed. by G. Meggle), Dr. Hänsel-Hohenhausen AG, Frankfurt: 459-474.

We-attitudes and Social Institutions

Petri Ylikoski & Pekka Mäkelä
Department of Philosophy
P.O. Box 24
00014 University of Helsinki
Finland

In this paper we want to examine the role of we-attitudes in the existing philosophical analyses of social reality. It is our view that we-attitudes have gained too central a place in the leading philosophical accounts of social ontology. (Searle 1995; Tuomela 1995, Tuomela & Balzer 1999) We acknowledge the importance of we-attitudes, but think their role in the analysis of social institutions has been exaggerated.

Unquestionably Tuomela and Searle have been the two leading pioneers of philosophical study of social reality and the nature of social institutions. There are many interesting philosophical questions concerning social institutions. As social institutions are weightless, non-extensional, invisible, tasteless, but still located in time and space, their ontology raises some vexing questions. Institutions seem to be causally efficacious in our social life, but how are these causal powers to be understood? What are the relations between the social and the mental, and between the social and the physical? Some of the questions have connections to epistemology: can there be institutions that nobody knows about? How is the relation between a fact and knowing a fact to be understood, if the fact is partly constituted by our knowledge of it?

The philosophical discussion has not yet provided answers to all of these questions. However, with respect to some issues a wide consensus prevails. For example, most writers agree that institutions are *performatively* created and that they have a nature of *collective good*. It is also agreed that social institutions include a component of *self-reference*, although this notion is un-

derstood differently by different authors. (See e.g., Searle 1995, Kusch 1999 and Tuomela & Balzer 1999.)

In this paper we will not discuss these broad issues. Our more modest aim is to concentrate on the role of we-attitudes in the philosophical analyses of social institutions. It seems that, at least among philosophers, there is consensus that we-attitudes have a central place in the philosophy of social institutions. Our main claims in this paper are that this emphasis on we-attitudes is not well-motivated and that philosophical theories of institutional reality should not restrict themselves to the analyses in terms of we-attitudes. In what follows, we will first study the positions of John Searle and Raimo Tuomela, who both are explicit supporters of the position we want to discuss. We will show how they make use of we-attitudes in their analyses and argue that there are more differences between these two authors than is usually acknowledged. In the second part of the paper we will sketch an argument to the effect that ‘strong’ we-attitudes are not necessary for the existence of social institutions.

1. The current consensus: analysis in terms of we-attitudes

It seems that most writers in the field of social ontology agree that social institutions are created by we-attitudes, or at least this thesis has not been challenged. However, this consensus is not very substantial. There is much variance in the accounts of we-attitudes, and sometimes we-attitudes are just taken as primitives. We will take John Searle’s and Raimo Tuomela’s analyses to be representative examples of this variety. We believe that if we can show that both accounts fail as analyses of social institutions, we can plausibly claim that we-attitudes are not a necessary element of social institutions.

According to Searle the basic formula for the creation of institutional reality is the following: ‘We accept (S has power (S does A))’. Roughly, this formula says that ‘we accept that items S count as A’. The acceptance by ‘us’ is a *performative act* which confers the *status* of A to items S. S’s being able to A is a *constitutive rule*, in the sense that it not only regulates how S can be used, but it creates or constitutes a new characteristic of S. The new characteristic is of the kind

that the item S does not have in virtue of its physical properties. Here the acceptance is to be understood in the broad sense, for example agreement, acknowledgement, recognition, going along with, etc. can amount to acceptance. Finally, the 'we' in Searle's formula refers to *collective intentionality*, which we will discuss below. The main idea is that by applying, modifying and iterating this basic formula we can reconstruct the whole ontological structure of institutional reality. (Searle 1995: 104-111)

We think that Searle's analysis is on the right track. The concepts of status function, performative act and constitutive rule will have an essential role in the ontological account of institutional reality. Our issue is with the role and status of collective intentionality in Searle's formula. To see the problem more clearly, let us see what Searle tells us about it.

Searle says that by collective intentionality he refers to the capacity to share intentional states, such as beliefs, desires and intentions, and to act on them. This capacity is a biologically primitive phenomenon for humans and some other animals, and it is not reducible to individual intentions. He does not think that the existence of collective intentionality requires the existence of collective consciousness, or collectives irreducible to individual agents. All that is needed for the existence of collective intentionality is that the members of the collective have appropriate we-attitudes. (Searle 1995: 23-26) Searle (1990) claims that we-attitudes cannot be reduced to individual attitudes, not even with the help of mutual beliefs. However, Searle's arguments to this effect are not very detailed and it seems that he misunderstands Tuomela & Miller's (1988) position that he takes to be a representative of a reductionist position.

Searle's theory does not require that agents understand the nature of social reality in accordance with his analysis. According to him, "... the participants need not be consciously aware of the form of the collective intentionality by which they are imposing functions on objects" (Searle 1995: 47) He allows, for example, that agents may accept imposition of a function on the basis of a theory that is not true. They may believe that their currency is "backed by gold" when it is not, or that their king's authority is divinely authorized. Similarly, agents need not understand relevant constitutive rules and concepts in the same sense as the outside analyst of the institution

explicates them. For example, when discussing constitutive rules concerning money, Searle writes: “The users of money do not know those rules, and in general [...] they do not apply them consciously or unconsciously” (Searle 1995: 142). Agents need not know, and generally they do not know, that they have we-attitudes.

When we look how the sentence quoted in the previous paragraph ends we gain greater insight into Searle’s position. He says, “... rather they have developed a set of dispositions that are sensitive and responsive to the specific content of those rules.” (Searle 1995: 142) Here Searle is referring to his concept of *Background* (Searle 1995: Chapter 6). This concept allows him to be very liberal about explicit beliefs of the agents and to avoid the idea of unconscious rule-following. The motivation for being liberal about explicit beliefs is easy to see. If we were to require that all agents have an understanding of social institutions that is in accordance with our theory, the applicability of the theory would be very limited. It is evident that most people have a very sketchy understanding of the social rules that they follow. We also agree with Searle that explaining rule-following by unconscious rule-following only adds to the mystery rather than provides any real illumination.

We fully recognize the importance of the idea of Background in principle. However, we find Searle’s analysis problematic. Our complaint is not with the idea itself, but with the sparse information Searle provides to support it. We are told that the idea is compatible with connectionist models of cognition (Searle 1995: 140-141), which is fine, but it does not help us much. It seems that in the end the whole idea boils down to the thesis that agents have behavioral dispositions and abilities that are tuned to the rules of the institution in such a way that agents behave *as if* they were following those rules consciously. And this is all that Searle provides us with.

We now come to the main problem in Searle’s analysis: *the whole idea of collective intentionality is just a black box*. He takes we-attitudes to be primitive, and this makes it unclear what it is to attribute an explicit we-attitude to an agent. One can accept that agents can have such attitudes, but how can one demarcate between having and not having a we-attitude? We are told that animals are also capable of collective intentionality. This does not help much in understanding

the role of we-attitudes in the constitution of social institutions. Unfortunately, Searle also has a very limited number of examples. He illustrates his idea by contrasting genuine cooperation with individual behavior that merely happens to be coordinated with the behavior of others. Collective intentionality is only present in the former. This example, and the earlier reference to the presence of collective intentionality in animals, suggest that Searle's we-attitudes refer to some basic form of sociality. This sociality would include responsiveness to others' actions and intentions, the capability to cooperate and so on. This sociality certainly includes the specific intentional attitudes that Searle discusses in his 1990 paper, but on the whole the content of this category is extremely heterogeneous. This interpretation would make the notion of we-attitude extremely weak. We think that this kind of sociality is not in contradiction even with a reductionist view of we-attitudes. Searle faces here a dilemma: either his thesis is a mysterious black box that does not provide us with much understanding or the idea is simply trivial.

The situation will not get better when we move to the 'as if' attitudes that are attributed to the Background. What is it to have an 'as if' we-attitude? Remember that we are basically speaking about behavioral dispositions and abilities, which are always properties of an individual. This would suggest that appropriate habits and capacities could be analyzed in strictly individualistic terms. And there are further problems. Consider the possibility that agents can have explicit beliefs that are in contradiction with the Background we-beliefs Searle attributes to them. This is possible because Searle allows that social agents need not know that institutions are constituted by their own beliefs. This seems to permit that agents' explicit beliefs include components that deny the claims made by Searle's analysis. So, in these cases we have agents with 'as if' attitudes that are in contradiction with their explicit beliefs. Is this kind of idea in any way coherent or sensible? All this erodes the whole idea of the Background we-attitudes. Bringing in the Background does not help Searle in his dilemma between apparent mystery and triviality.

By contrast with Searle, Tuomela provides us with a very detailed analysis of we-attitudes. According to Tuomela, the idea of collective acceptance (CA) is intended to be very generally applicable. What is this collective acceptance that can work wonders?

The CA model of sociality concerns the collective acceptance of sentences ("ideas") for the use of the group. The model says that the very acceptance of a sentence creates its correct assertability for the group in question. It is meant to be a general account of collective-social "things". Arguably it covers all social institutions. A social institution in the broad sense, according to Tuomela, is a type of recurrent norm-based collective activity, that is, a type of recurrent collective activity in which some forms of sanction or pressure are present.

Tuomela accepts, in accordance with the recent philosophical and sociological literature (Barnes 1988, Kusch 1999, and Searle 1995), that many social (or collective-social as Tuomela says) things and their characteristics are performatively created, and that some central social concepts are reflexive. Tuomela's account adds a third feature of sociality, the collective availability or *forgroupness* of collective social items. The forgroupness idea is meant to capture the indexicality of the social institutions. Social entities always exist in relation to some specific social group or collective. However, Tuomela's forgroupness involves more than just the indexicality, since it is closely related to the notion of collective acceptance.

Tuomela's analytical apparatus consists of the idea of acceptance of sentences, the notion of sentential attitude, as well as the notions of correct assertability and truth. Collective acceptance covers both the production and the maintenance of "social things", and Tuomela speaks about notions and sentences. Collective acceptance amounts to coming to (collectively) hold, and holding certain collective thoughts or ideas (viz. we-intentions and or we-beliefs), and being disposed to act on those ideas. The core idea of Tuomela's account is formulated in his *collective acceptance thesis* (CAT), where *s* may be an arbitrarily complex sentence:

a sentence *s* is *collective-social* [...] in a group *G* if and only if (a) it is true for group *G* that the members of group *G* collectively accept *s*, and that (b) they collectively accept *s* if and only if *s* is correctly assertable (or true). (Tuomela & Balzer 1999: 181)

In Tuomela's account collective acceptance in the sense of coming to hold an idea for a group is collective social action in the sense of an action performed for a shared social reason. Social rea-

sons are we-attitudes in the following technical sense. (We-attitudes encompass, e.g. we-intentions, we-wants, we-beliefs)

#) A has a we-attitude to X in group G if and only if A shares X in the mode of that attitude and believes that X is so shared in G and believes also that it is mutually believed in G that X is so shared in G. (Tuomela 1995: 38)

For example, an agent we-believes that p in G if and only if he

believes that p

believes that all members in G believe that p

believes that it is mutually believed in G that all the members in G believe that p.

We-attitudes can be either in I-mode or in we-mode. In the CA account the we-mode is required.

In CAT clause (a) expresses the assumption of the categorical collective acceptance of s, while clause (b) characterizes the kind of collective acceptance which is needed here. Tuomela's distinction between I-mode and we-mode is of importance here, as he argues that "the intentions and beliefs have to be in the we-mode (or group-mode) rather than in the I-mode in contexts of collective acceptance for the group" (Tuomela & Balzer 1999: 179). As Tuomela understands it, "the we-mode involves 'we-thoughts' in the sense of group members being collectively committed to seeing to it that what is collectively accepted is made correctly assertable or regarded as correctly assertable by group members for the group members." (Tuomela & Balzer 1999: 178)

We-attitudes in the I-mode in contrast express a weaker kind of we-ness, as there is no joint or collective commitment, or forgroupness involved.

In CAT collective acceptance is assumed to entail forgroupness and collective commitment to the accepted sentence. "... when accepting something for the group the participants are collectively committed to a rule system which in general requires that the members perform certain actions (e.g. inferences) and permits the performance of some other actions." (Tuomela & Balzer 1999: 180-181)

In the light of these considerations, and with some plausible assumptions, the analysans in CAT can be reformulated in the following way:

a sentence s is collective social in a group G if and only if (a) the members of group G collectively accept s , and (b) they collectively accept s if and only if s is correctly assertable for group G .

CAT is the basis for Tuomela's account of social institutions. In his 1999 paper we find the following general characterization of social institutions:

(SI) A norm-entailing sentence s expresses a social institution in a primary sense in a collective G if and only if the members of G collectively accept s for the use of G , with the understanding that collective acceptance for the group entails and is entailed by the correct assertability (or truth) of s for G . (Tuomela & Balzer 1999: 198; For an alternative formulation, see Tuomela 2000: 175.)

Let us see how this analysis works in Tuomela's favorite example of an institution: "squirrel fur is money". In this case the s must express something like "squirrel fur counts as money". Now, Tuomela must mean that the agents do not just collectively accept this sentence for use in their discourse, but also in their actions. That is, the agents must *use* squirrel furs as money. The analysis (SI) presupposes that the agents understand that their collective acceptance of squirrel fur as money makes it money for them. This expresses Tuomela's requirement that the members of G understand 'the true nature of collective acceptance' which is a necessary condition of social institutions in his account. This is quite strong a requirement. From our point of view, the requirement that the agents understand the true ontological nature of their institutions is too far-fetched. The agents do not need to understand either the indexicality or the performative character of social institutions, not even dispositionally. These requirements have not always been fulfilled in the history of human institutions. After all, people can have false beliefs about the ontological nature of their institutions.

Our problem with this analysis is that it is too demanding. It seems to us that requiring collective acceptance in the we-mode requires too much. We will provide arguments for this thesis in the next section. Before going on to these arguments, let us make a note about the aims of Tuomela's analysis of social institutions. Tuomela claims that his analysis applies to all 'collective-social

things', but it is not clear that he is analyzing actual social institutions. From some passages one gains the impression that he is accounting for ideally justified institutions. In such case, perhaps, the we-modedness requirement can be successfully defended. For example, Tuomela writes: "Even if in actual life collective acceptance may be collectively non-intentional, it could have been collectively intentional; in this our account resembles political contract theories" (Tuomela & Balzer 1999: 179). This interpretation is supported by passages in *Cooperation* as well. There Tuomela speaks about "justified institutional action" (Tuomela 2000: 173). In the same passage he accepts "institutional acting without collective commitment" as a possibility (Tuomela 2000: 173). If this is true, we do not have much criticism against his account in this paper. However, we believe that Tuomela is also trying to account for the existence and maintenance of all real-world institutions, and in that case we think that the conditions given by Tuomela are not all necessary for the maintenance of social institutions.

2. Institutions without we-attitudes

Both Tuomela and Searle have originally developed their ideas of we-attitudes in the context of collective action. (Searle 1990; Tuomela 1984, 1995, 2000) We do not have much to say about these analyses in this context. We are not claiming, for instance, that collective intentions can be reduced to individual intentions. Our concern is with the use of we-attitudes in the analysis of institutional reality. Let us assume that we-attitudes are essential in the analysis of some forms of social action. Does this suffice to show that all social institutions, as they are normally understood, should be analyzed in terms of we-attitudes? Clearly not – some further arguments are required. Neither Searle nor Tuomela has provided an explicit argument for *this* thesis. We also claim that such arguments should not be expected to be successful.

Our specific thesis is that we-attitudes in the we-mode are not *necessary* for the existence of social institutions. We are not saying that they are not *sufficient*. Our examples are not intended to be examples of typical institutions, but examples of realistically possible institutions. This is all that is required for our argument. We will not discuss here the issue of the creation or the emer-

gence of institutions. Our concern is the ‘maintenance’ problem, the question concerning what is required to sustain institutional reality and to keep the institutions going. We will follow Searle and Tuomela and use an elementary form of money as our example of a social institution. The following cases are, for brevity’s sake, ‘snapshots’ of participants’ beliefs about some *ongoing* institutions. So we assume that coins are actually used as a means of exchange. The beliefs are not just theoretical speculations of the participants, but premises of their practical reasoning.

C1: All members of community C believe that the monetary value of their coins is a natural property of coins.

It seems that in this case we can have a perfectly working institution. As all agents believe that the value of their coins is a natural property, monetary exchange can continue as smoothly as in any other case. The use of money in this case requires as much (we-mode) we-attitudes as using a screwdriver. The only requirement is that a sufficient number of participants have these beliefs, and that they act on the basis of them. It is the *use* of the coins (with certain value) by individuals that constitutes them as money. The institution can continue its existence as participants teach the appropriate beliefs to newcomers. And as everything works fine they do not have any reason to doubt their beliefs about the nature of monetary value.

Of course, the agents probably believe that *they* (the members of C) are using money. In monetary exchange (weak) we-beliefs are plausibly involved. However, the important point is that these we-beliefs do not *constitute* the institution in question. To see this point consider an example of another elementary social institution, authority.

A1: All members of community C believe that the authority of their leader X, is his natural property.

Now, in this example we can see the point about the ontological constitution of social institution more clearly than in the case of money. In contrast with monetary exchange in this example, the members of C do not need to have any beliefs about each other at all. It is sufficient that everybody individually believes that X has authority of a leader and acts on this belief. No mutual be-

liefs among the members of C are required for a perfectly functioning institution. The mutual beliefs, or we-beliefs, are not necessary for the existence of a social institution in this case.

These cases are direct counter-examples to Tuomela's analysis. Consider the case of coins. The agents do not have the required we-attitudes. Similarly, the agents do not believe that the value of their coins is only for their group and constituted by them, as Tuomela requires. The only way for Tuomela to save his analysis, is to make a similar move to the 'as if' analysis as Searle does. So, let us see how Searle's analysis succeeds with this example.

Searle allows that situations like **C1** are possible, so this is not a direct counter-example to his theory. He would say that agents have these explicit beliefs and a set of 'as if' we-attitudes in the Background. However, the example raises the question whether Searle's we-attitudes are doing any work in cases like this. *If we can have an institution which runs without we-attitudes, why include them in the analysis at all?* Searle clearly needs to give an account of the import of we-attitudes in this case. It seems to us that everything we-attitudes are supposed to contribute can also be achieved without them. We take **C1** to be either a clear counter-example to Tuomela's and Searle's analyses or at least a serious pointer for the need of explication of the philosophical import of the we-attitudes to the analysis of social institutions.

We can add to the difficulties of the 'as if' analysis by modifying the example in the following way:

C1*: All members of community C believe that the monetary value of their coins is a natural property of coins and they explicitly deny all suggestions that the coins have value because people believe them to have such a value.

In this case, we have exactly the problem that was mentioned earlier: the explicit beliefs of the agent deny the 'as if' beliefs that are attributed to the agent. Because 'as if' beliefs are not real beliefs, not even unconscious beliefs, there is no formal contradiction here. However, the idea that they are somehow analogous to the actual beliefs is in jeopardy. At least, an informative account of the nature of 'as if' beliefs is needed.

Let us take a different example.

C2: All members of community C (only) believe that other members of their community believe that the monetary value of coins is a natural property of coins.

In the previous example the institution was constituted by purely individual beliefs. In this case it is constituted by the members' belief that others have such beliefs. So, instead of 'we' or 'I', we have 'others', but still we would expect that the institution in question can run smoothly. Note that agents in this case might lack the personal belief that the monetary value of coins is their natural property. In this case, every individual agent believes that if others believe that coins have a certain value, then they have such a value. And if the agent believes that everybody else has such beliefs, then he has good reasons to act as if he believed that the value of coins is a natural property and to use them. It does not matter that no-one actually believes that the value of coins is a natural property. Everything is based on what others are believed to believe. As a consequence, it does not matter if we change the example a little:

C2*: All members of community C believe that other members of their community believe that coins have certain monetary value because they believe that everybody else believes so also.

If **C2** works, and our analysis of it is correct, **C2*** will also work. Both **C2** and **C2*** work without we-attitudes. So, basically it is possible that all agents in this community are pure opportunists. There is no personal or collective *acceptance or commitment* in such cases. At most individuals go along with the idea. And there is no way of turning this individual form of opportunism into the we-attitudes that Tuomela or Searle require without trivializing their analyses.

3. Discussion

Someone might ask why we have considered as complex situations as **C1**, **C1***, **C2** or **C2*** when the following simple example would do: everybody just uses coins as a medium of exchange. In principle, this is correct. If we were to offer a positive proposal about the constitution of the social institutions, it would be in this ballpark. However, we chose these examples here in order to provide a slightly stronger argument, that would also work against an 'as if' strategy.

Our thesis in this paper is that the maintenance of social institutions is not necessarily based on we-attitudes in either Searle's or Tuomela's strong (we-mode) sense. Institutions can easily work without them. However, someone might suggest that there are some roles for we-attitudes that we have not taken into account. For example, it might be suggested that the normative elements in the institutions require the presence of we-attitudes. We have two replies. First, it is not clear that all institutions have such elements. Second, it is not clear how the we-attitudes account for the normative elements at all. However, the issues related to normativity are a topic for another paper.

Bibliography

Barnes, Barry 1988: *The Nature of Power*. Polity Press. Cambridge.

Kusch, Martin 1999: *Psychological Knowledge. A social history and philosophy*. Routledge. London.

Searle, John 1990: 'Collective Intentionality and Action', in Cohen, Morgan & Pollack (eds.) *Intentions in Communications*, MIT Press. Bradford: 401-415.

Searle, John 1995: *The Construction of Social Reality*. Allen Lane, Penguin Books. London.

Tuomela, Raimo 1984: *A Theory of Social Action*. D. Reidel. Dordrecht.

Tuomela, R. & Miller, K. 1988: 'We-intentions', *Philosophical Studies* 53: 367-389.

Tuomela, Raimo 1995: *The Importance of Us*. Stanford University Press. Stanford.

Tuomela, Raimo 2000: *Cooperation*. Kluwer Academic Publishers. Dordrecht.

Tuomela, R. & Balzer, W. 1999: 'Collective Acceptance and Collective Social Notions', *Synthese* 117: 175-205.