

Chapter 12

Consensus

Learning from Max Weber's Problem

Max Weber's status in social science in general, and in German social theory in particular, is ambivalent. On the one hand, his importance as one of the foremost classics of social theory and social science is uncontested. On the other hand, however, he is routinely accused of relying on a skewed methodology. Especially, Weber's claims concerning the action theoretic foundations of sociology have been criticized. Because Weber's action theory is at the very heart of his work, and cannot be separated from his sociological theory, this is no insignificant charge.

In the development of German social theory, particularly during the last decades of the twentieth century, this diagnosis has played a crucial role. During this time, the field was divided into two camps, and the interpretation of Weber's work served as one of the battlegrounds for their controversies. The first camp's label was sociological Systems Theory, with the later Talcott Parsons and Niklas Luhmann as its main protagonists. The second camp was gathered around the label "Critical Theory", with Jürgen Habermas as its theoretician-in-chief. Both camps claimed to be able to provide a solution to the problem of Weber's action theory, while accusing the other camp of failing miserably at this task.

Much ink has been spilled on the controversy between the two camps, and the very peculiar role Weber seemed to have played in this has not gone unnoticed, either. Yet it has not been noticed thus far that there is one particular Weberian category that plays a crucial role in this. It is a category that, even in the enormous body of literature on, Weber does not seem to have received the attention it deserves: the category of *Einverständnis* and *Einverständnishandeln*, which Weber analyzes in the sixth section of his essay *On Some Categories of Interpretive Sociology* which was published in 1913.¹

Einverständnis and *Einverständnishandeln* are no easy terms to translate. In the existing translations, the former is usually rendered either as "agreement" or as "consensus", while the term "consensual action" seems to be the preferred choice as far as *Einverständnishandeln* is concerned. As always, each of these translations has its problems. "Agreement" seems to have the advantage of being more

¹ Weber ([1913] 1981). German original: "Über einige Kategorien der vertsehenden Soziologie" (1913) in Weber (1922): *Gesammelte Aufsätze zur Wissenschaftslehre*, hereafter quoted as WL.

common in ordinary English language than “consensus”, but in order not to sever the conceptual link to “consensual action”, I shall either use “consensus”, or the German original in the following.²

The term *Einverständnis* plays a dual role in Weber’s action theory. On the one hand, it stands for a basic phenomenon of social reality. On the other hand, it indicates a fundamental problem for Weber’s action theory. To put it in Weberian terms (which will be explained below), the *phenomenon* is this: there is a special case of social action in which the action, in its “subjectively intended meaning”, is oriented towards other agents’ actions in the form of *normative expectations*. Now why should that phenomenon be a problem for Weber’s theory? It is this: how can this phenomenon (and with it the role of social normativity in general) be accommodated within an action theoretical framework that assigns a paradigmatic role to instrumental rationality and goal-oriented action? Or, to put the problem more simply: why should rational instrumental agents ever *have* such expectations?

Parsons, Luhmann, and Habermas unanimously believe that the problem cannot be solved, because the phenomenon simply goes beyond the conceptual capacities of Weber’s (or, in Habermas’ diagnosis: the *official* Weber’s) action theoretical framework. And both camps – Systems Theory as well as Habermas’ version of Critical Theory – draw their consequences by departing from Weberian action theory and by basing their respective theoretical edifices on new foundations. In Parsons’ and Luhmann’s case, this is done by switching from an action theoretic framework to systems theory. Habermas, for his part, disavows Weber’s instrumentalism and intentionalism, and turns to a linguistic account that assigns the paradigmatic role to communicative action rather than to goal-oriented action.

Both camps claim to have solved Weber’s problem. Interestingly, however, they deny the other party any recognition of the achievement they claim for themselves. Thus, in Luhmann’s view, Habermas simply remains stuck in the old action theoretic conceptual framework within which it is simply impossible to deal with the problem, whereas Habermas makes no secret of his view that Luhmann, with his functionalism, fails miserably to overcome the mentalistic instrumentalism of Weber’s theoretical framework.

Given this constellation, and the importance it had for the development of social theory, it might be worth the effort to take a closer look at the source of all this trouble. What precisely is the phenomenon, and why exactly does it not seem to fit into Weber’s theoretic framework? And, above all: how can Weber’s problem, if there is any, be solved: is it the Habermasian or rather the Luhmannian solution that works – or neither of the two? And if so: is there a third way?

² Wherever possible, my translations will follow Richard Swedborg’s *Max Weber Dictionary* (2005).

§40 The Problem of Interaction

In his essay *On Some Categories*, as in his other writings, Weber claims that the proper object of interpretive social science is a particular kind of action. In Weber's account, action is defined as intended behavior, i.e. behavior that has some "subjectively intended meaning". For an action to be the proper object of interpretive social science, it has to meet the following requirements:

Action significant for interpretive sociology is (...) behavior that (1) in terms of the subjectively intended meaning of the actor, is related to the *behavior of others*, (2) is *codetermined* in its course through this relatedness, and thus (3) can be intelligibly *explained* in terms of this (subjectively) intended meaning (Weber 1981: 152).

Weber labels this kind of behavior "*Gemeinschaftshandeln*", a term which he uses elsewhere in his work, and which is usually translated as *communal action*. In the only existing translation of the essay *On Some Categories*, however, the term is rendered with "social action", which also translates Weber's *soziales Handeln*. What is the reason for this? First, it is obvious from Weber's definition that, in the essay *On Some Categories*, the term *Gemeinschaftshandeln* is used in the very same sense as the term *soziales Handeln* has in the rest of Weber's work, and especially in Weber's *magnum opus*, which was posthumously edited under the title *Economy and Society*. Weber also uses the term *Gemeinschaftshandeln* in *Economy and Society*, but he does so in quite a different meaning. *Gemeinschaftshandeln* is here defined as a small subclass of *soziales Handeln* (a type of action which is marked by the fact that the agents are tightly connected by a sense of belonging together).³

This latter meaning, i.e. the use of the term *Gemeinschaftshandeln* in *Economy and Society*, seems to correspond much better to the intuitive notion of "communal action." As we shall see, the meaning in which Weber uses the term *communal action* in his essay *On Some Categories* might appear highly counterintuitive. But still, we should not ignore that he does use the term here, and we should, I think, resist the temptation to correct Weber's terminology. Especially since the following interpretation is largely based on Weber's essay *On Some Categories*, I will use Weber's terms as he defines them in that essay. I will translate *Gemeinschaftshandeln* literally, i.e. as "communal action", but it is important to remember that this term is equivalent to "social action" as defined in *Economy and Society*. I will not talk about the narrow definition Weber gives the term "communal action" in *Economy and Society*, so I will only use that term in the sense of Weber's earlier essay.

Why does that term seem so strange that even the translator of the essay chooses to replace it with "social action"? The reason is simple. One of the most obvious characteristics of communal action, as defined in *On Some Categories*, is that it does *not* presuppose any community. Therefore, communal action simply seems to be a *misnomer*. As quoted above, communal action (in the sense of the essay

³ In *Economy and Society*, *Gemeinschaftshandeln* is as a type of primary group action "based on the subjective feeling of the parties, whether affectual or traditional, that they belong together". Cf. Graber (1981).

On Some Categories) simply requires that some agent's behavior be oriented, in its subjectively intended meaning, towards another agent's behavior. It is true that it takes at least two for an action to be communal – there has to be at least one other agent around, to whose behavior the agent is oriented – but it seems that there does not have to be any community, or, in Weber's terminology, any form of *social relation* between the two whatsoever. Assume that A is the agent, and B is the other person towards whose behavior the “subjectively intended meaning” of A's behavior is oriented. For there to be communal action (or, in the terminology of *Economy and Society*, social action) as defined by Weber, it is not necessary that B in any sense *participate* in the matter. Indeed, B does not even have to *know* what the “subjectively intended meaning” of A's behavior is, or indeed that there is another agent. This is what makes talk of *communal action* particularly counterintuitive. It is true that the examples Weber uses are usually of a different and more communal character, but if we stick to the letter of the definitions Weber gives us, there need not be any reciprocity whatsoever for there to be communal action. B might think he is all alone on the planet; yet, as long as A, who is creeping up behind him to mug him, effectively orients the meaning of his behavior to B's, his action is “communal action”, even though there is nothing communal about that action in the intuitive sense of the term.

As far as I know, there is no clear indication to be found anywhere in Weber's work that Weber ever thought this definition of the primary object of interpretive social science to be fundamentally deficient. But there are some hints that he thought that some specifications were in order. Thus, in the third paragraph of *Economy and Society*, Weber focuses on what he calls “social ‘relation’”. In the case of social relations, the orientation towards the other agent's behavior is *reciprocal*. Social relation is defined as the “actions of several persons that are mutually adjusted and oriented to each other in their meaning (. . .). A minimal degree of relation of *reciprocal* orientation (. . .) is an essential trait of the concept”. Most certainly, it is an advantage of Weber's concept of social relation, rather than any shortcoming, that it does not favor *cooperative* kinds of social relations over *competitive* kinds. “Enmity” and “competition” figure prominently on Weber's list of social relations, among many cooperative examples. Yet there is something else that should be noted, and that should be carefully distinguished from the fact that Weber leaves room for conflict-laden forms of social relations. Weber also allows for the possibility that, for all the mutuality and reciprocity of the subjectively intended meaning, social relations might nevertheless be “objectively ‘one-sided’”, as he puts it. This is the case where the agents, as Weber continues, “attach different meanings to their behavior.” Weber's further explanations make the reader think that Weber, still under the label of “social relation”, even wants to allow for cases in which there is mutual misunderstanding, i.e. in which the agents mutually ascribe attitudes to each other which they do not actually have. This is just another effect of Weber's basic action theoretic commitment, according to which the “subjectively intended meaning” that is constitutive of action need not be “true” or “correct” at all. The consequence for the concept of social relation is this: any trace of actual reciprocity is stripped off that concept; objective reciprocity is, as Weber says, simply a “limiting case” of social relation. Thus it seems even social relations need not in any sense be communal.

Weber thinks that social relations are simply a special case of the meaning structure of social action. Upon closer consideration, however, it seems that the step from social action to social relations is a decisive one for the entire project – at least as far as Weber’s claim that “interpretive sociology (...) is not part of psychology” (Weber 1981: 154) is to be taken seriously. The step is from the analysis of single agents’ actions to the “behavior of a plurality” (*Sichverhalten mehrerer*), as Weber says in §3 of *Economy and Society*. The question to be addressed in the following is this: can this transition really succeed within Weber’s action theoretical approach? As it shall turn out, the category of *consensus* plays the key role in this transition.

But first, it has to be shown why this transition should be a problem for Weber at all. In order to do so from an *immanent* perspective, it is necessary to give a somewhat broader account of Weber’s theory of the structure and role of subjectively intended meaning in paradigmatic cases of action. We have already introduced the terms “communal action” (or *social action* in the terminology of *Economy and Society*) and *social relation*. Let me briefly introduce the most important Weberian distinctions of action types (where not otherwise indicated, the following quotations are from the first paragraphs of *Economy and Society*).

Clearly, the type and structure of subjectively intended meaning can vary widely, depending on the kind of social action in question, just as long as there is some orientation towards another agent’s behavior. First of all, action need not be *rational*. Weber distinguishes two kinds of non-rational action. Non-rational action can be either *traditional* action, i.e. action that follows some “blind routine” (*dumpe Gewöhnung*), or it can be of the *affective* kind. In contrast to traditional action, affective action involves a clear consciousness of the action. Social action, too, can be non-rational, because routine behavior and emotional acts can be oriented towards other agent’s behavior. Just as important as the distinction between rational and non-rational forms of action is Weber’s distinction between two *types of rationality*. First, there is *value rationality*. Value-rational behavior is “determined by a conscious belief in the intrinsic value of some ethical, aesthetic, religious, or other form of behavior, independent of its prospects of success.” Thus value-rational action shares with affective behavior the element of clear consciousness, which is absent in traditional action. By contrast to both forms of non-rational action, however, value-rational action does involve systematic planning (this is what makes value-rational behavior *rational*). The fourth and last action type is *instrumental rationality* (*Zweckrationalität*). “Instrumentally rational behavior is behavior exclusively oriented to means (subjectively) considered adequate to attain goals (subjectively) clearly comprehended” (Weber 1981: 151; Weber’s definition of instrumental rationality in *Economy and Society* will be discussed below). Like as value-rational action, instrumental rationality involves consciousness and planning. In contrast to value-rationality, however, instrumentally rational agents do not disregard their prospects of success. Rather, the effects and side-effects (i.e. both the intended and the unintended consequences) are taken into account.

Thus Weber’s action theory allows for a considerable variety of action types. Yet there is another side to the matter. If it is true that there is much space for different kinds of action, it is also true that Weber’s methodology clearly favors *one* type,

which serves as the paramount case, or the paradigm of action.⁴ There is one type of action which is the *fullest* type, and which therefore is action *par excellence*. Concerning the distinction between rational and non-rational forms of behavior, the emphasis is clearly on the former – in spite of Weber’s occasional claims to the contrary. The reason for this rationalistic bias is partly methodological. Weber’s rationalism (a label which he rejects) is most evident where Weber uses rationality to distinguish the realm of objects and events that can only be *explained* (i.e. the sphere of natural science) from the realm of the proper objects and events for *interpretation* (e.g. WL 67). Such remarks make one think that Weber is not really serious after all about the possibility of genuinely non-rational action (for a certain behavior would not count as an action if it were plainly non-rational), but even where Weber explicitly does allow for irrationality in the domain of action, he clearly favors rational forms of action over irrational ones – both at the methodological level, and at the level of the content of the analysis. It is true that we can “understand even the irrational exertions of the most excessive emotions”, but if Weber claims that according to his view, these are “just as accessible” to interpretation as “the chains of rational ‘reasoning’” (WL 100), he seems to be slightly exaggerating, to say the least. Weber’s methodology clearly favors rational action, because rational behavior is simply more intelligible than non-rational behavior. And within the domain of rational action, it is *instrumental* rationality that “possesses the highest measure of ‘self-evidence’” (Weber 1981: 151). Instrumentally rational action simply has the “most understandable kind of meaning structure” (WL 408; 127, a statement which clearly flies into the face of any claim to a non-rationalist hermeneutics). Thus it is no coincidence that instrumental rationality is the *default* action type that is presupposed in Weber’s famous methodological tool, the *ideal type*.

It appears that both the critical and the defensive view of Weber’s action theory have their point. It is true that Weber does allow for a rich variety of action types. But it is also true that he methodologically narrows his focus on one particular kind of action, namely action of the instrumentally rational kind. On this line, it seems that Weber believes instrumentally rational action as the “fullest” action type, the other types being more (traditional and affective action) or less (value-rational action) deficient.

The primacy of instrumental rationality over value rationality is particularly evident from Weber’s remarks on communal action in his essay *On Some Categories*. Weber does allow for the possibility of value-rational communal action. But the instrumentally rational case is just as paradigmatic for the structure of meaning of communal action as it is for the meaning of any action. Communal action that is motivated in “ideal values”, and that is oriented to norms of conduct or to the fulfillment of duties without there being any calculation of the consequences involved, does occur, but Weber considers this as just a “limiting case”. In the “normal case”, Weber says explicitly, communal action implies a degree of *instrumental*

⁴ It is not surprising that, in the received literature, these two tendencies separate Weber’s critics from his defenders. For the latter see, e.g., Baurmann (1996: 283ff.).

orientation.⁵ To the degree that communal action is instrumentally rational, the agent does not simply do what she thinks she *owes* to the other (in terms of duty, or in terms of any other ideal values such as good taste). Rather, the decisive element of communal action is the following. The agent takes the other agent's *expected behavior* into account in calculating the best means to realize her goals.

An important [...] normal component of communal action is its meaningful orientation to the *expectations* of certain behavior on the part of others and, in accordance with that, orientation to the (subjectively) assessed probabilities for the success of one's own action. (1981: 159)

This is parallel to the case of non-communal (or, in the terminology of *Economy and Society*, non-social) action. Here, the agent takes into account the expected restrictions and the given circumstances of the objective surrounding world while calculating the best means to his or her end. The only difference between non-communal and communal instrumentally rational action is that, in the latter case, the circumstances that have to be taken into account happen to be other agents' behavior. Significantly, Weber's definition of instrumental rationality at the beginning of *Economy and Society* covers both cases. A certain behavior is instrumentally rational if it is determined by the agents' "expectations of the behavior of objects of the external world and of other human beings", with "these expectations serving as 'conditions' or 'means' for rationally pursued, weighed goals" ([1921] 1980: 12; hereafter quoted as WG). Instrumentally rational communal agents simply do whatever is best to realize their goals (weighing the effects against the expected side-effects), *given the expected behavior of the other agents*. The basic structure of expectation is, as Weber says explicitly, "basically (...) the same" whether B is an inanimate object, or whether he is another agent.

At the same time, however, Weber sees that there is one decisive difference involved here, which captures nicely the modern distinction between paradigmatic and strategic rationality. Let me quote the decisive passage from Weber's *On Some Categories*:

A subjectively rational agent can also base his expectations of certain behavior from the part of the others on his subjective belief that he can expect subjectively *meaningful* behavior from others, and that he can thus predict, with varying degrees of accuracy, the probabilities arising from certain meaning relationships. (1981: 159)

The obvious difference between the case in which B is a stone and the case in which B is another agent is this: in the latter case, the latter behavior is *action*, i.e. it has *meaning*. Is this just one case among many as the passage seems to suggest? Or is there more to that fact, so that the expectation of B's behavior is not "basically [...] the same", after all, but of an entirely different kind? Let's have a closer look at the matter.

⁵ This is particularly obvious from the fact that, in the later chapters of his essay *On Some Categories*, Weber seems to regard a feature that is only characteristic of instrumentally rational communal action to be a feature of all of communal action, namely the (cognitive) expectation of the other's behavior as a base for the calculation of one's own course of action.

First of all, it seems plausible to ascribe to this case a much more prominent a role for the concepts of social action or indeed communal action than Weber does. If we strictly stick to the letter of Weber's definition in *Economy and Society*, it does not seem necessary for A to conceive of B as an agent for his action to be social (or communal in the terminology of *On Some Categories*), just as long as B *is in fact* an agent. But this does not seem plausible. There is no reason why A, who mistakes B for a stone, and whose action, in its subjectively intended meaning, is oriented to B's behavior, should thereby be a case of social action. The orientation to another agent's behavior that is part of the definition of social action cannot be simply *de facto*; rather, A has to *believe* that B is an agent, just as is implied in the above quotation. (Another question, which will not be addressed here, is whether *belief* is sufficient, or whether *knowledge* is required. What about an agent who mistakes a stone for a person? Is this a case of social action?)

Let's call this the action-orientation condition. I think it is plausible to assume that it has to hold where action is social. For action to be social (or communal), the agent has to orient his or her behavior not just on some other agent's *behavior* (which he might mistake for a natural event). Rather, he or she has to see that the behavior in question is *another agent's*, and that it is *subjectively meaningful* for him or her (i.e. the other agent). In other words, the agent has to take the behavior in question to be *an action*. It seems quite obvious to me that the action-orientation condition conforms to the spirit of Weber's theory, even though it might be the case that, in *Economy and Society* at least, it cannot be found there explicitly.

If this is the case, communal actions are cases in which A orients his behavior to B's expected behavior, which he interprets *as action*. If this is true, however, a problem pops up. To understand the issue at stake here, it is important to keep in mind Weber's abovementioned distinctions, i.e. the distinction between rational and non-rational behavior, the distinction between value-rational and instrumentally rational behavior, and, above all, the distinction between "one-sided" or unilateral social action on the one hand, and social action of the kind of *social relation* on the other. Let's start with the paradigmatic case of social action in Weber's theory: unilateral instrumentally rational action. Here, there seems to be no problem with the above condition that the orientation to another agent's behavior cannot simply be *de facto*. A, who hides behind a tree to ambush B does not expect B's coming by the tree to be a mere natural event (as he expects the apples to fall from the trees in autumn), but an *action*, i.e. as motivated in B's putative aim to go for a walk, or some such. It seems clear that an A who meets the action-orientation condition has certain advantages over an A who sees B's behavior merely as a natural event. The advantage concerns what instrumental rationality is all about: the prospects of success. If A perceives B's behavior *as action*, and grasps the *subjectively intended meaning* of the behavior in question, he will be able to form more reliable predictions and be more successful in forming the corresponding expectations concerning B's behavior. Thus, for the case of unilateral instrumentally rational action, action-orientation is a simple instrumental advantage in terms of goal effectiveness.

What about the case of social relation? What about *reciprocal* instrumentally rational action? If (1) A bases his choice of means on his expectations concerning

B's behavior, and (2) A believes that B is instrumentally rational, too, and if (3) the orientation is (assumed to be) *mutual*, the situation in question meets the criteria of Talcott Parson's concept of *interaction*:

In interaction ego and alter are each objects of orientation for the other. The basic differences from orientation to nonsocial objects are two. First, since the outcome of ego's action (e.g. success in the attainment of a goal) is contingent on alter's reaction to what ego does, ego becomes oriented not only to alter's probable overt behavior but also to what ego interprets to be alter's expectations relative to ego's behavior, since ego expects that alter's expectations will influence alter's behavior. Second, [...] this orientation to the expectations of the other is reciprocal or complementary. (Parsons and Shils 1959: 105)

Parsons (1951: 10) labels the implicit structure of *interdependent expectations* with the term *double contingency* (or *multiple contingency*). This structure is most important for the development of Parsons' thought. The idea of double contingency is conspicuously absent in Parsons' earlier work, especially in his *Structure of Social Action* ([1937] 1949) even though, in his interpretation of Weber's action theory, Parsons had all the necessary means to see this structure. It seems obvious that there is a connection between Parsons' becoming aware of the structure of interaction, and his later turn to systems theory. Indeed, his turn away from action theory and towards systems theory seems to be *motivated* by the view that it is simply impossible to fully accommodate the structure of interaction in an action theoretical conceptual framework (as we shall see, this line of reasoning is particularly obvious in the case of Luhmann's systems theory).

But why should double contingency be a problem for action theory – and an insoluble one at that? The following interpretation of the problem imposes itself. If A is instrumentally rational, he will have to base his decision over the choice of available means (behavior) on an expectation concerning B's behavior. *If* he expects B to be calling back, he should wait, if he doesn't, A should call B himself. This, however leads into an *infinite regress* where A believes that B is instrumentally rational, too, and meets the action-orientation condition. A now believes that B bases her choice of means (behavior) on his (B's) expectation of A's behavior. The expectations are, in other words, *mutually interdependent*, and cannot serve as a basis for the choice of means. The issue at stake here is currently being discussed in Rational Choice Theory, where the focus is on the question of how agents can coordinate their choices rationally.⁶ In Weber's terms, the problem is that, in such situations, A's orientation to B's "subjectively meaningful" behavior does not lead to an expectation of probable courses of actions on which A can base his choice of means. If according to Weber's definition, instrumental rationality requires A to base his choice of means on his expectation concerning B's choice, this presupposes that A's expectation and his choice are independent from each other. This, however, is not the case where A assumes that B's choice will be based on B's expectation concerning A's choice, and where A assumes that there is a mutual belief that this is the case (this need not be *common knowledge*). Thus it is obvious that, to the agents,

⁶ Cf. Schmid (2007), Chapter 6 in this volume.

this double contingency (or multiple contingency) poses an insoluble “decondition-alizing problem.” There is simply no firm ground on which instrumentally rational agents can base their calculations of the best means to realize their goals. Instead of leading to better predictions, and more accurate expectations (as in the one-sided case mentioned above), agency-orientation here leads instrumentally rational agents into total paralysis. Instead of coming up with any prediction of B’s likely behavior, A gets lost in an infinite regress, forming an expectation of B’s expectations of A’s expectations of B’s expectations, etc. pp. The question is: how can this consequence be avoided? How can, in other words, the problem of strategic interdependence be solved?

§41 Consensus

In all of Weber’s methodological work, there seems to be no proof that he ever became aware of the full extent of the problem of interdependent expectation and its consequence for his theory of instrumental rationality. Indeed, Weber has often been criticized for not addressing the problem of strategic interdependence of decisions at all.⁷ In the parlance of today, Weber’s theory of rationality in action got stuck on the level of *parametric rationality*.⁸ But there are some passages that can easily be read as indicating some degree of awareness of the issue at stake here. The most striking example is a statement in the essay *On Some Categories*. Weber here talks about the role of expectations in instrumentally rational forms of communal actions, which leads him very close to the problem of strategic rationality. Weber says here of the expectations characteristic of communal action that they are marked by an “absolute instability” (WL 422). Significantly, he continues by saying that communal action needs to be *normatively supported* or *integrated* in order to be stabilized. Talking of the agent, he continues, almost hastily:

specifically, his expectations may be based on an ‘understanding’ with another or with others; he then believes that he has reason to expect compliance with the ‘agreement,’ according to the meaning which he himself attributes to it. This alone is enough to give communal action a specific qualitative particularity, for this significantly enlarges the area of expectations toward which the actor believes he can rationally orient his actions. (Weber 1981: 159–160)

In the light of the abovementioned problem of the strategic interdependence of expectations, this *element of normativity* seems to be *mandatory* for rational interaction to be possible at all, and not just one option among others. If this is true, Weber’s claim seems to have far-reaching consequences for the architecture of his theory. If it is true that mutual, instrumentally rational communal action (i.e. an instrumentally rational “social relation”) *presupposes* some form of explicit mutual

⁷ For a very clear statement of this fact see Norkus (2001).

⁸ “The parametrically rational actor treats his environment as constant, whereas the strategically rational actor takes account of the fact that the environment is made up of other actors, and that he is part of their environment, and that they know this, etc.” (Elster 1979: 19).

commitment to a normative order, or some form of explicit agreement, to which the agent expects the relevant others to stick, there simply is *no conceptual room left* for this particular action type. For instrumentally rational social action based on explicit agreement comes very close to – and indeed is *identical* with – another one of Weber’s categories. It is *associational action* (*Gesellschaftshandeln*). In *On Some Categories*, Weber gives the following definition of the term:

Communal action shall be called ‘associational action’ when and insofar as (1) it is oriented in meaning toward expectations that are held on the basis of agreements, (2) the formation of these agreements has resulted purely rationally (*zweckrational*) in view of the expected action of the associated persons, and (3) the orientation of meaning is subjectively rational. (Weber 1981: 160)

The problem is this. We have seen earlier that, for Weber, the paradigmatic case of action is instrumentally rational action, that the proper object of social science is communal action, and that the paradigmatic case of communal action is reciprocal. Now, if this is what *associational* action is, Weber’s claim that associational action is simply a *special case* of communal action, and that there are other forms of communal action, even paradigmatic ones, that are not of the associational kind, seems to be undermined. To uphold this architecture of his theory, Weber needs to identify cases of communal action that are reciprocal, but not of the associational kind. And indeed this seems to be not only a question of the structure of his theoretical edifice, but required by the “things themselves”. Reciprocal communal action *is* more fundamental than associational action, and indeed associational action *presupposes* that there are communal actions of the reciprocal kind. How else if not by means of some reciprocal communal actions should the formation of agreements on which associational action is based ever come about? If there are no social relations, and if the relations are not reciprocal, there cannot be such a thing as an agreement, and therefore no agreed upon social order. Without acting in reciprocal social relations, people cannot enter an agreement, and form a contract, and cannot, therefore, perform associational actions (which presuppose agreements, contracts, or some such social orders). But the extension of the concepts of “agreement” and “reciprocal social relations” are not simply co-extensive. All agreements imply reciprocal social relations (I am in an actual agreement with you precisely insofar as you are in an actual agreement with me), but not all reciprocal social relations are agreements of the explicit contractarian kind that is presupposed in Weber’s concept of associational action. Rather, such agreements are reciprocal relations *of a special (and especially complex) kind*. Therefore, it would be a mistake to approach the structure of reciprocal social relations from the analysis of the structure of agreements. The analysis should run the other way around: we need to understand the structure of reciprocal social relations first, before any progress can be made in the analysis of the structure of agreement.

If this is true, we should hold on to the overall architectural plan of Weber’s action theory, at least as far as the relation between communal action and associational action is concerned. Associational action is a subclass of communal action. But if the foregoing conjecture is true, if we need a concept of social *normativity* to understand how communal action can be reciprocal, and if that normativity cannot

be pulled out of the hat of agreement (because agreement presupposes reciprocal communal actions), we are left with the following question: what kind of action can there be that is normatively stabilized (thereby avoiding the problem of the “absolute instability” of interdependent expectations), yet not be based on agreement?

This is the point where we finally come to the topic of this chapter. For this is precisely the role of *consensus* and *consensual action* in Weber’s theory. *Consensus* is the missing link between communal action and associational action.⁹ The term *Einverständnis* (consensus) describes the fact that agents can form reciprocal expectations without getting lost in some infinite circle or loop of interdependent expectations, because agents can be linked by some form of *mutual commitment*, without there being agreement or some explicit normative social order involved. It is characteristic for Weber’s concept of consensus, and indeed one of its advantages, that *consensus is not based on agreement*. Weber shows clear awareness of the fact that consensus, if the term is to play its structural role in the architecture of the theory, cannot be understood as some “tacit agreement”, either. The reason is simply that consensus is whatever makes agreements, including tacit ones, normatively binding (cf. WL 433), and not the other way around. Thus consensus is the source of all social normativity, of institutionalized forms (contracts) as well as of other forms. Weber defines “consensus” in the following words:

[‘Consensus’ is] the fact that an action oriented on expectations concerning the behavior of others has an empirically realistic chance of seeing these expectations fulfilled because of the objective probability that these others will, in reality, treat those expectations as meaningful and ‘valid’ for their behavior, despite the absence of an explicit agreement. It is conceptually immaterial which motives underlie these expectations about this behavior of others. Communal action insofar as it is oriented on such probabilities of ‘consensus’ shall be called ‘consensual action.’ (Weber 1981: 186, and also WL 432)

Against the background of the above considerations, this definition raises a whole series of questions. First and foremost, the normative element needs to be questioned. At first sight it seems that the element of “validity” concerns B’s expectations only, and does not entangle A. If there is consensus, B takes A’s expectation to be of some normatively binding quality to him. Yet Weber continues by emphasizing that the reason why A should expect B to feel somehow normatively bound by A’s expectation are “immaterial”, thereby suggesting that the point is not that B really has to treat A’s expectations as “valid” for there to be consensus, but that A has to *believe* that B does so. For there to be consensus, it is not enough that (1) A expects B to behave in a certain way, and that (2) B treats A’s expectation as normatively binding. Rather, (3) A has to *believe that B treats his expectation as normatively*

⁹ It is true that, in his essay, *On Some Categories*, the section devoted to *consensus* (which fills almost half of the entire essay!) comes only after the section on associational action (which in turn follows the analysis of communal action). But it is clear that, within the logic of Weber’s action theory, consensual action occupies the middle position. Indeed this is the succession of terms when Weber enumerates the types of action elsewhere (cf., e.g., WG 381): “communal action, consensual action, and associational action”. This mirrors the foundational structure: association presupposes consensus, and consensus, in turn, presupposes communal action.

binding (for whatever reasons A believes B to have for this) and that B conforms to A's expectation "in reality" *because* he does so. This not only seems to be Weber's view; it is also more plausible, for it excludes the case in which A expects B to conform to A's expectations for any other reasons. Consider the case in which A expects B to do x because A thinks x maximizes B's pleasure, while B knows that A expects him to do x and does x because he believes that he *should* conform to A's expectations. This should not be considered a case of consensus for two reasons. First, it does not seem plausible to call any of A's action that are based on such expectations *consensual* because this is counterintuitive given the ordinary language meaning of the word. Second, and much more importantly, such cases do not explain how the problem of interdependent expectations can be solved, because for this to be the case the normative element has to be part of *both agents'* "subjectively intended meaning".

If it is only *accidental* that B conforms to A's expectation for normative reasons, and not in any way *expected* by A, it seems plausible to call his action *social*, but not in any way *consensual*. For pure social action, it is immaterial what motifs agents ascribe to each other. If, however, the general aim of Weber's methodology is to construe the essential categories of action theory in terms of *subjectively intended meaning*, Weber cannot use "objective probabilities" of behavior to distinguish social action from consensual action; rather, the *differentia specifica* between consensual action and other kinds of social action has to be a matter of the *content* of the agent's "subjectively intended meaning", i.e. of what he intends and what he believes. Thus consensual action presupposes that A takes B to fulfill A's expectations for the reason that B thinks that he *should* do so. In the reciprocal case, he will not only have to take B to assume that he, A, is normatively bound, too. Above all, he has to take *himself* to be normatively bound to conform to B's expectations.

It seems that the kind of expectations at stake here defines the concept of a social norm. According to the simplest definition, a social norm is a special type of social regularity. A social regularity is a reason to expect that people will exhibit a certain type of behavior. If the regularity in question is of the special kind of a social norm, there is a reason to believe that people will exhibit the type of behavior in question *because they believe they can (justifiably) be expected to do so*.

If this reading of the structure of consensual action is right, it seems, however, that there is still something deeply wrong about the architecture of Weber's theory. The problem is that the characteristic feature of consensual action seems to be more than just a simple *differentia specifica* of social or communal action.

Upon closer consideration, it appears that Weber's talk of expectations in the definition of consensual action quoted above, and the use of the word in the definition of social or communal action, covers up an ambiguity in that term. The difference at stake here is between expectations of the *cognitive* kind, on the one hand, and *normative* expectations, on the other. While cognitive expectations are part and parcel of the concept of social action (as the foremost way in which agents take into account other agent's behavior), it might seem that the expectations in the above definitions are of the *normative kind*. Thus the question whether or not consensual action fits into the general theory of social action depends on the relation between these two kinds of expectations.

Here are some differences between the two. Most obviously, the *direction of fit* is different. In the case of cognitive expectations, the direction of fit is mind-to-world; in the case of normative expectations, it is world-to-mind. The difference between the two directions of fit becomes particularly obvious in the case in which the expectations are not fulfilled. If A *cognitively* expects B to perform some action, and if it turns out that B does not perform the respective action, A will put the blame on himself rather than on B. A sees now that he simply *miscalculated* B's behavior, i.e. that his expectation was *mistaken*, which is his problem, not B's. If, however, A *normatively* expects B to perform an action which he or she does not perform, A will put the blame on B rather than on himself: it's not that A was *mistaken* in expecting B's action; rather, *it is B's fault not to do what he or she was expected to do*. In the first (i.e. the cognitive) case, A shouldn't have expected B's action; in the latter (i.e. the normative) case, B shouldn't have failed to do what was expected from him or her.

This is the reason why the latter type of expectation is generally much more resistant to contradictions with experience than the first. Cognitive expectators are ready to *learn*, i.e. to adapt the content of their expectations to what they know about the structure of the world (hence the mind-to-world direction of fit). One cannot cognitively expect one's pet to be housebroken if one believes it is not. But one can *normatively* expect one's pet to be housebroken if one thinks that this is what the pet *should* be. *Normative* expectators are more likely to *teach* than to learn, to change the world rather than their expectations, i.e. to see to it that the empirically observed behavior in question meets their requirements (hence the world-to-mind direction of fit).

Another way to approach this difference between cognitive and normative expectations focuses on a difference in the kind of *intersubjective relations* involved in each of these cases. The relation between the subject of the expectation, and the person whose behavior is expected is fundamentally different: cognitive expectators *take other people's behavior into account*, while normative expectators *count on other people*. This difference is fundamental indeed, and it is obscured not only by the ambiguity of such words as "expectation", but of other words such as the verb "to rely on", too. If we climb up a big tree, we rely on the strength of the branches in a sense that is fundamentally different from the sense in which we rely on our comrade while climbing on a rope in the mountains, and it may well be that our last words will mark this difference if our expectations should be disappointed.

The question to be answered here is the following. If the expectations involved in consensual action are of the normative rather than of the cognitive kind, does that mean that Weber can assign this type of coordination its place within the framework of his theory only at the price of an equivocal terminology? Does Weber use the double meaning of the word "expectation" for the purpose of pulling consensus out of the hat of social action?

This is precisely what both of the opposing camps in German social theory seem to think. In the following, we shall have a closer look at how Niklas Luhmann and Jürgen Habermas try to solve Weber's problem, and how their choice of basic conceptual framework is determined by their attempt to succeed at this task. Parsons

and Luhmann think that this cannot be done within an action theoretic framework, and that it is necessary to turn to systems theory in order to get a grip on the phenomenon. Habermas, in turn, sticks to action theory, but switches from intentionalism to a linguistic foundation for social theory. I believe that neither Luhmann's nor Habermas' solution really works, and I will try to show why. In the last section, I shall make a suggestion as to how Weber's problem could be solved within action theory, and within an intentionalist setting at that. All that is needed is a theory of *collective* intentionality.

§42 Consensus and Contingency

For the later Parsons, the problem of double contingency and the role of social norms in interaction ultimately show that social theory cannot be based on an account of (rational) action. The action theoretic focus on "subjectively intended meaning" is simply too narrow to capture the structure of interaction, because within this framework, it cannot be explained how agents come to have normative expectations. Consensual actions imply some kind of normative order: that normative order is necessary to solve the problem of the circle of interdependent expectations, and to counteract the "absolute instability" (Weber) of strategic interdependence.¹⁰ In other words, action theory cannot explain, but indeed *presupposes* the existence of a normative order, at least if it extends to actions of the consensual kind. Since there is no doubt that interaction is in fact possible, and since this cannot be explained within an action theoretic framework, the structure of interaction cannot be captured within an account that is focused on the "subjectively intended meaning" of behavior. Interaction is not a matter of isolated individual *unit acts* (Parsons' term). The actions which interacting agents perform have to be taken as elements *within a system*.¹¹ Interaction requires a consensual element, and that, in turn, presupposes the existence of a generalized normative order, i.e. a shared system of symbolic meaning, conventions, and other cultural standards. Because these cannot be reduced to the "subjectively intended meaning" of the participating individual agents, the theory of social action has to go beyond action theory. Consensus cannot be understood from the perspective of the subjectively intended meaning of the agents; rather, the analysis has to account for the structural conditions of possibilities for consensual action. And this is precisely what systems theory is all about. This radical shift in

¹⁰ "The most important single condition of the integration of an interaction system is a shared basis of normative order. Because it must operate to control the disruptive potentialities (for the system of reference) of the autonomy of units [...] such a basis of order must be normative" (Parsons 1968: 439).

¹¹ This is already expressed in Parsons ([1937] 1949: 740). This is an early indication of the line of development James Coleman criticized in the following words: "Parsons abandoned his attempt to found social theory in a theory of rational action; he reverted to classification schemes that were no less sterile in his hands than in the hands of those he criticized" (Coleman 1992: 49). For Coleman's own treatment of the problem of double contingency cf. Coleman (1990: 901ff.).

perspective entails a complete reversal of the order of analysis. Instead of giving an account of social facts based on the interpretation of the subjectively intended meaning of the agents (as it is done in action theory), social facts are now, in turn, seen as based in the “mutuality of socially structured relationship patterns,” (Parsons 1954: 359) and these are taken to be the *preconditions for the possibility* of such individual mental states, which have to be listed and classified in order to develop an adequate account of interaction. Thus Parsons’ use of the argument that social norms are a condition for the possibility of interaction is such that the action theoretical, hermeneutic approach is replaced by an analytical, classificatory reconstruction of systemic structures from an external point of view.

To the critics, this radical shift of perspective clearly shows that systems theory does not *solve* the problem, but simply discards the point of view from which it is a problem. The perspective is shifted from the search for *reasons for action* that are fit candidates to rationalize normative expectations to the search for the structural preconditions for empirical motivational patterns. Of course, Parsons’ underlying thesis is that the structure of the meaning of consensual action simply goes beyond the subjective intention, i.e. beyond the perspective of the participating agents, and is, therefore, of systemic nature. According to Parsons, the cultural symbols (which include normative components) are responsible for the success of coordination, and thus have to be taken into account even though they are more like *causes* than *reasons* for consensual action. Loosely speaking, it is the system that is in control of the situation, not the agents themselves, “the system being so *geared into the action system* of both ego and alter that the external symbols *bring forth* the same or a complementary pattern of orientation in both of them. Such a system of normative orientation is logically the most elementary form of culture” (Parsons et al. 1959: 16; my emphasis). As Parsons continues, this “internalization of culture patterns” *creates* “personality” (ibid.: 22) as part of the social system, which in turn means that Parsons conceives of personality entirely in terms of cultural conventions. In a sense, the system replaces the individual as *the agent* of consensual action. This might seem implausible in itself, yet there is another consequence which seems particularly hard to swallow. The view that social agents are “steered” by social norms is at the heart of what has rightly been criticized as Parsons’ “oversocialized” account of action, leading to Parsons’ notorious conventionalism. From the point of view of systems theory, *all* action seems to be norm-regulated behavior, with norm being somehow the more fundamental category than action in the order of the concepts. The phenomena of deviance and dissidence, conflict and innovation are excluded from the analysis on the conceptual level (cf. Wrong 1992: esp. p. 216).

This is an immediate effect of the way in which Parsons tries to solve the problem of Weber’s theory by reformulating the normative component of consensus as “normative social order” in a systems theoretic perspective. And even if we leave aside the point that Parsons’ systems theory *avoids* Weber’s problem rather than *solving* it, and that the way in which the system replaces the agency as the center of control seems somewhat questionable, these consequences seem much too difficult to accept calling Parson’s way of dealing with Weber’s problem successful.

It is interesting to see how Niklas Luhmann, in his version of systems theory, addresses this problem. Luhmann tries to overcome Parsons' conventionalism, which is a consequence of Parsons' turn to systems theory, with his own version of systems theory. The third chapter of *Social Systems*, Luhmann's *opus magnum*, is entitled *double contingency* (Luhmann 1984: 148ff.). It starts with a critical discussion of Parsons. Luhmann here argues that Parsons' mistake was to introduce his concept of normative order as some kind of *compensation* for the problem of double contingency, i.e. as some external element that makes double contingency *disappear*, instead of seeing double contingency as a conceptual ingredient, and indeed as a precondition for the possibility of any communication. In other words, Luhmann seems to think that Parsons' "oversocialized" view is caused by the half-hearted way in which he dealt with double contingency, and by the quest for intersubjective "convergence" of meaning. With a more courageous shift to systems theory, Luhmann claims, all of that is left behind. Consensus is a problem only for those who think that there really *is* such a thing as intersubjective identity of meaning. By renewing systems theory on the base of the concept of *autopoiesis*, Luhmann claims that he can finally liberate social theory of all older claims to consensus and intersubjective convergence. The mistake of all older theories, Luhmann claims, is to start out with the category of meaning, which is relative to a subject or system, but then to claim that different subjects and systems can somehow converge, and to construct meaning as an inter-subjective category. According to Luhmann, this attempt is simply futile, and should be discontinued. If meaning is relative to one system of subject – as Luhmann, too, thinks it is – it is relative, full stop. Thus the theory of autopoietic systems is aimed at taking the ontological subjectivity of meaning seriously. The modern experience that meaning can vary, that there is no inter-subjective or inter-systemic unity that guarantees some prestabilized harmony, some convergence, let alone some universal identity, is not taken as a sign of crisis which has to be overcome, but as a fundamental feature of social reality which we should finally take for what it is: everything could always be different. There is no place for sure bets in social reality, not even in what might appear to be the firmest consensus. Thus in his reading of double contingency, Luhmann stresses the modal feature of the word contingency:

Contingency means that being depends on selection which, in turn, implies the possibility of not being and the being of other possibilities. A fact is contingent when seen as selection from other possibilities which remain in some sense possibilities despite a selection. (Luhmann 1976: p. 509)

The fact that, in the situation of strategic interdependence, individuals cannot base their expectations on each other's choices, i.e. that the attempt to base one's own choice on the other's expected choice leads into an infinite regress of interdependent expectations, does not mean, according to Luhmann, that some external normative element (normative social order) has to be there so that interaction can start. The fact that individuals are "black boxes" to each other, the "darkness of mutual intransparency", as Luhmann puts it in this chapter, is nothing that has to disappear for communication to become possible. The whole point of Luhmann's theory is to drop the idea of the convergence and mutual transparency of the meaning of

behavior. Mutual opacity and the relativity, indeed, the *privacy*, of meaning is not a problem which has to be overcome for the social system to be possible; rather, these are *preconditions for the possibility of the social system*. The social system does not emerge *in spite*, but *because* of double contingency. The element of commonality which Parsons declared to be a condition for social action, and which he found in the normative social order, is simply declared inessential and consequently dropped. The social does not have to be anchored within the realm of the intended meaning of the participating individuals. In Luhmann's view, the individuals are – and remain – utterly opaque to each other. The social does not resolve double contingency; rather, it emerges from *whatever happens* in situations of double contingency. This is what communication is, according to Luhmann: whatever happens in situations of double contingency. Thus double contingency defines and underlies the social. And because the social is based in double contingency, it is not to be reduced to the participating individuals (individual systems), but rather an emergent level of system.¹²

For a social system [...] it is not necessary that the systems which are in double contingency to each other can see through each other, and prognosticate. The social system is a system for the very reason that there is no basic certainties of state, and no predictions relying on these. (Luhmann 1984: 157)

Luhmann defines the elements of the social system as *communication*. Communication, just like the expectations of the participating individuals, is a matter of *meaning*. But, as is apparent from the above considerations, the term “meaning” is rather equivocal in Luhmann's theory. Meaning in terms of communication is not meaning in terms of the intentions of the participating individuals. Quite to the contrary, Luhmann's way of dealing with the concept of double contingency implies the radical distinction, the categorical rift between the two kinds of meaning, for which his theory is justifiably infamous: the distinction between communication (i.e. meaning at the level of the social system) and consciousness (i.e. meaning at the level of the thoughts of the participating “psychic systems”). Luhmann rejects any attempt to go from meaning in terms of whatever is “subjectively intended” to meaning in terms of communication. In a kind of a bold strike through the Gordian Knot of double contingency, the meaning of meaning is simply cut in half: meaning

¹² The decisive passage in *Social Systems* reads as follows: “The black boxes generate whiteness, as it were, when they meet, or at least sufficient transparency to deal with each other. By means of *their mere hypothesizing* they generate certainty of reality, because that hypothesizing leads to the hypothesis of hypothesizing from the part of *alter ego*. The assimilation of meaning materials to this level of order presupposes [...] two self-referential systems, which observe each other. For those few aspects which are of importance for their dealings, their capacity for information processing might be sufficient. They remain separate, they do not merge, they do not understand each other better than before. They concentrate on what they can observe about the other *qua* ‘system-in-an-environment’ in terms of input and output, and they learn self-referentially, within their own observer perspective. They can try to influence whatever they observe by means of their own action, and they can learn from the feedback they get. In this way, an emergent order can come into existence, an order that is dependent on the complexity of the systems that make it possible, but that does not depend on the possibility that this complexity can be calculated, or controlled. We call this emergent order the social system” (Luhmann 1984: 156f.).

is the medium of the integration of psychic systems as well as of social systems. But “social meaning” (in terms of communication) cannot be derived from “psychic meaning” (in terms of the meaning of the thoughts of individuals), and neither can the content of thoughts be derived from communication. Thus communication is not shared meaning, but something entirely different.

From the point of view of the architecture of the theory, this move might seem fascinating. From an analytical perspective, however, this is so utterly implausible that it seems difficult to understand why Luhmann’s systems theory has been successful for so long. Indeed, the distinction seems rather self-defeating. Why should we *think* there could be anything to the things which Luhmann *tells* us if there is no bridge between what is told and what is thought?

Granted, the two do not always converge. In an introduction to Luhmann’s systems theory that is quite widely used in the German speaking world, the consequences of Luhmann’s categorical rift between thought and communication is illustrated with the example of a physician’s wandering thoughts during an interview with his patient. In cases of routine conversation, such phenomena as talking on auto pilot might be quite frequent indeed, but it is simply nonsensical to declare this the normal case. Even if they are wrong more often than they think, people say what they say because they assume they say what they think, i.e. that the meaning of communication reflects the meaning of the speakers’ thoughts.

This is very basic folk psychology and common sense indeed, and to go against it would require better reasons than finding an elegant solution to the problem of double contingency. Not surprisingly for a social theorist, Luhmann himself is not willing to stick to his own claim, when his concern is not with the conceptual aspects of his grand theory but with its descriptive content. More than anywhere in the gargantuan body of Luhmann’s work, this is the case in his writings on education science. Analyzing the relation between the student’s thoughts and the communication going on in the classroom, Luhmann departs from his idea of an unbridgeable difference between thought and communication, and instead claims that there is some “congruence (...) of psychic and social events” (Luhmann 1987: 179). Luhmann does not explain any further what he means by “congruence”, but it seems quite clear that he does not do so because, whatever congruence might turn out to be, it will hardly be compatible with his clear conceptual cut between thought and communication.

Thus Luhmann’s own analysis proves his “solution” to the initial problem wrong. We are still stuck with the task of finding a way to show how interaction and “subjectively intended meaning” relate to each other, i.e. of what reasons agents might have for normatively expecting other agent’s behavior. Luhmann’s bold stroke through the Gordian Knot does not make the problem disappear. The trouble with Weber’s category of consensus remains. Many of the complexes of behavior that are of interest to social science are based on the fact that individuals do count on each other, and can successfully do so. In these cases, the “expectations concerning the behavior of others has an empirically realistic chance of seeing these expectations fulfilled because of the objective probability that these others will, in reality, treat those expectations as meaningful and ‘valid’ for their behavior, despite the absence of an

explicit agreement” (Weber 1981: 168; WL 432). An essential feature of these cases is the following. By contrast to Luhmann’s claim, these agents do not see each other simply as “black boxes”, or as “opaque systems”, and do not limit themselves to observe their input and their output. The relation between such individuals (or systems) is of an entirely different stamp indeed. These individuals (or “psychic systems”, if you wish) are related to each other by *normative* expectations; they do not just *take each other into account* as elements of the surrounding world; they *count on each other*. This phenomenon (which Weber labels “consensus”) cannot be accounted for within a conception that dichotomizes the sphere of meaning into “psychic” and “social”. Thoughts and communications are different in many respects. But their relation is more intricate than Luhmann believes. In trying to avoid the problematic consequences of Parsons’ attempt to do justice to the role of social normativity in interaction, Luhmann lets himself be carried away into a conception that not only seems rather self-defeating and to fly into the face of common sense. Above all, his conception ends up in a straight denial of the phenomenon which Weber’s category of consensus is supposed to capture. In his view, systems always “remain separate, they do not merge, they do not understand each other better than before. They concentrate on what they can observe about the other *qua* ‘system-in-an-environment’ in terms of input and output, and they learn self-referentially, within their own observer perspective. They can try to influence whatever they observe by means of their own action, and they can learn from the feedback they get.” (Luhmann 1984: 156f.)

In other words, systems (psychic or other) do not have normative expectations concerning other system’s behavior, because all they do is observe their input and their output, which – if there are any regularities to be observed – leads them to cognitive expectations. But, in reality, there *are* normative expectations. To learn more about their nature, role, and structure, Luhmann’s systems theory is obviously the wrong place to turn.

§43 Consensus and Language

Let’s now have a look at the other of the two camps in which German social theory has been divided over the last 2 decades of the twentieth century, and see how Weber’s problem is dealt with there. For Jürgen Habermas, Weber’s essay *On Some Categories* proves that there is an *unofficial version* of Weber’s action theory, a version centered on the concept of *consensus*, hidden behind the official version. The official version is, according to Habermas, marked by a concept of meaning which is oriented towards the paradigm of the solitary instrumental action of a single subject. This concept of meaning stands for everything Habermas thinks we have to leave behind in order to proceed to an adequate conceptual framework for social theory. First, the official concept of meaning is *intentionalistic* instead of *linguistic*: it captures meaning in terms of whatever is “subjectively intended” by the agent, and tries to derive the structure of consensual action from the “subjectively intended

meaning” of the agents. Second, Weber’s concept of action is *instrumentalist* rather than *norm-oriented*. On the “official” line, Weber describes action primarily in terms of means and goals, i.e. as a matter of the pursuit of instrumental success rather than in terms of social propriety, conformity, or deviance. And third, Weber’s account of action is *monological* rather than *communicative*. Action appears as a matter of single individuals who act in light of their own goals, rather than as a matter of group members who act on a shared understanding of their situation.

Thus his “official” Weber nicely epitomizes Habermas’ opponents: *Bewußtseinsphilosophie*, intentionalism, functionalism, “monologism”. Against all of that, Habermas recommends no less than a radical shift of paradigm. This is achieved with his new “linguistic foundation of sociology.” Meaning, Habermas claims, should be seen as a matter of language rather than as a matter of intentional mental states. Following a great many philosophers of his generation, Habermas argues that linguistic meaning cannot be derived from the intentionality of the mental states of the speakers. Rather, the intentionality of mental states is *derived* from linguistic meaning. Thus, in his view, meaning is no matter of “subjective intention”, as Weber has it, and is thus not only *contingently* social (insofar as others are in the content of these mental states). Rather, meaning is a matter of shared linguistic practices, and hence *a priori* social (cf. Habermas 2001).

Later on in Habermas’ work, however, it becomes clear that Weber plays a *dual role* within the project of the *Theory of Communicative Action*. On the one hand, Weber still serves as the epitome of everything that has to be left behind. On the other hand, however, Weber’s essay *On Some Categories* serves Habermas as evidence for his claim that, deep inside the “official” Weber, there is a little “unofficial” Weber struggling to get out. Needless to say that this unofficial Weber is much more to Habermas’ taste, since he almost seems to preempt Habermas’ own shift of paradigm towards a linguistic foundation of social theory. Following is what Habermas says about Weber’s “unofficial” views.

Whereas the “official” Weber tries to conceive of human coordination mostly in terms of the agents’ *interests* (*Koordination durch Interessenlage*), the author of the essay *On Some Categories* contrasts this mechanism with an entirely different type of coordination: *coordination by consensus* (*Koordination durch Einverständnis*). It is not difficult to recognize that the distinction between these two types of coordination anticipates Habermas’ later fundamental distinctions between *System* and *Life-world*, or between instrumental and communicative action.

In Habermas’ reading of the unofficial Weber, the distinction between communal and associational action is *within* the domain of coordination by consensus. These action types are, according to Habermas, sub-classes of consensual action: communal action is consensual action with a low degree of rationalization, while associational action is highly rationalized, with rationalization meaning value-rationalization rather than instrumental rationalization. Habermas then reads Lawrence Kohlberg’s dichotomy of conventional and post-conventional moral reasoning right into that distinction. The result is the following. In the case of communal action, agents can count on other agent’s behavior insofar as their expectations conform to the conventional standards of social propriety. In the case of

associational action, the normative expectations of the participating agents are based on the formal principles of fairness.

As a consequence of Habermas' reading, consensus is conceptually cut off from the sphere of instrumental action. Thus this reading is in a sharp contrast to another interpretation that seems to be imposed by the general architectural structure of Weber's theory, i.e. the attempt to find the sources of consensus and of the social normativity which is presupposed in consensual action somewhere *within* the domain of instrumental action. In the larger context of Habermas' general venture, the *Theory of Communicative Action*, this point is by no means a marginal one. Immediately following (and indeed based on) this somewhat idiosyncratic reading of Weber's categories, Habermas introduces the most important of his own conceptual tools, i.e. the distinction between *communicative* and *strategic action*:

We call an action [...] *strategic* if we consider it under the aspect of compliance to the rules of rational choice and if we assess the degree of efficacy of the influence on the decision of a rational opponent. By contrast, I speak of *communicative* actions, if the plans for action of the participating agents are coordinated *via* acts of communication rather than *via* ego-centric calculations of success. In communicative actions, the participants are not primarily oriented towards their own success; they pursue their individual goal under the condition that they can tune in their plans for action on each other on the base of shared definitions of the situation. (Habermas 1981: 385)

In brief, and put very simply, communicative action *is* Weber's consensual action. The special flair of Habermas' reading of the normative element in Weber's concept of consensual action stems from the fact that Habermas claims the element of communication within consensus to be *of a linguistic nature*. Consensus is ultimately the "telos of human language", and to be achieved *only* by linguistic means: "The concepts of speaking and of agreement (*Verständigung*) mutually interpret each other." (Habermas 1981: 387) Thus there is, according to Habermas, a necessary link between the meaning of consensual action and the linguistic competence of the respective agents. As Habermas puts it, a consensus is necessarily a proposition that is mutually accepted. Thus consensus has a "linguistic structure" (*ibid.*: 386).

I do not think that it is necessary to waste much time with the question of whether or not Habermas is right when he claims that his reading of the "unofficial version" of Weber's theory is "well supported" by Weber's essay *On Some Categories*, as far as interpretive correctness is concerned. It is simply too obvious that Weber does not have anything like Habermas' distinction in mind here, but aims at integrating consensual action into the theory of communal action in just the way as he does elsewhere in his work. In contrast to what Habermas claims, communal action is introduced as the *general term* that encompasses both consensual action and associational action. This makes it utterly implausible, from an interpretative standpoint, to open up a conceptual rift between goal-oriented action and consensus in Habermas' sense. But let's leave these tedious questions of interpretative correctness behind and turn to an issue that is of much greater relevance: is Habermas' concept of consensus *qua* theory of the source of social normativity consistent and analytically adequate?

The main problem I wish to address is the necessary link between speech and consensus as postulated by Habermas. This seems problematic for two reasons. First and foremost, this theory *a priori* limits consensual action to linguistic practitioners, which does not seem particularly convincing considering everyday examples in which consensual action seems to be oriented towards non-linguistic agents. And second, it seems that we need some concept of consensus to analyze the preconditions that enable pre-linguistic beings to enter linguistic communication. These beings have to *see each other* in a certain way, and it seems that the concept of commitment nicely captures some necessary features of their mutual relation. If consensus *is itself* a matter of speech, however, this concept is useless for the analysis of those preconditions of speech.

Let us have a closer look at Habermas' claim first. Of course, Habermas does not want to suggest that each and every consensus has to be a matter of explicit linguistic communication. There is what Habermas calls *lebensweltliches Vorverständnis* – a kind of consensus that is based in the meanings of the life-world, which agents simply take for granted in their every day dealings. But, in Habermas' view, the only reason why agents take those tacit life-worldly background assumptions to be consensual is that they *think that possible dissent could be negotiated by means of entering into a discourse, which is a linguistic matter*. In other words: if A tacitly assumes that B treats her expectation as valid for his own behavior, she does so because she thinks that, if B were to disagree about the validity of her expectation, they could have a discussion about this and find their way back to a consensus. Thus, in this sense, even those consensual meanings which agents do not talk about, and simply take for granted, are *constituted* by language (in terms of the possibility of a discussion). Consensual agents not only have to assume that their counterparts will accept their expectations as valid; they also have to see each other as possible candidates for an open debate, if the validity of the respective expectations is contested.

This thesis does not seem particularly appealing. Especially if we stick to Weber's definition, it does not seem obvious why consensual action should presuppose the possibility of linguistic communication. Indeed it seems quite frequent empirically that people have normative expectations concerning the behavior of beings, which they *do not* seem to take to be competent linguistic practitioners, or indeed possible candidates for discussions on contested claims to validity. Beings capable of speech are not the only members of the class of actual addressees of normative expectations. Among the beings of whom a certain behavior is normatively expected are those who are not linguistic practitioners *yet* (and perhaps some who are not linguistic practitioners *any more*). And indeed there are some members of this class, which never were – and never will be – competent speakers. This cannot be accommodated in Habermas' theory; without the possibility of linguistic communication, there simply cannot be normative expectations; the only attitude that is possible towards non-linguistic beings is cognitive expectation.

But how do we determine whether or not A expects B's behavior in a normative or in a cognitive way? Even though these kinds of expectation are very different conceptually, as the above considerations have shown, it is difficult to distinguish the two kinds empirically. This is especially so since A's behavior in reaction to

a disappointed expectation does not by itself tell whether his expectation was of the cognitive or of the normative kind. Let's assume that A is a dog owner and B is his pet. If A finds out upon his return home from work that his expectation concerning his pet's being housebroken has been disappointed once again, he may react by drawing the dog's attention to its pile, shout "fie!", and perhaps do some of the other things people usually do in such situations. The point is that from this behavior alone it is not possible to tell which of the following scenarios is true. According to one interpretation of the dog owner's behavior, he is *disappointed* to find out that his expectation was mistaken, because his dog is obviously not yet house-broken after all, and needs some more training. His behavior is an attempt to condition the dog so as not to exhibit that kind of behavior again. According to the other interpretation, however, the dog owner is genuinely *angry* at the dog, because he thinks that his dog did something it *shouldn't* have done, and that he was *justified* to expect house-broken behavior of his dog, and now *punishes* it for its deviant behavior.

These are two very different scenarios from the internal perspectives, even though from an observer's point of view, the behavior is exactly the same. In the first case, the disappointed expectations are of the purely cognitive kind, in the second of the normative kind.

Aside from the epistemological problem of how to determine, from a third person's perspective, if a given case of expectation is of the normative or of the cognitive kind, it does seem obvious that people *do in fact* address normative expectations to beings with whom they could not have a discussion. Of course, such phenomena are per se no counter-evidence to linguistic theories of social normativity such as Habermas'. The mere *fact* that some agent *assumes* herself to be in some kind of consensual relation with a being that is incapable of speech does not prove that non-linguistic consensus is possible. It may well be that consensus is *merely assumed* by the dog owner, which is not quite enough for there to be *actual* consensus. Even according to Weber's notoriously subjectivist definitions, there needs to be more than just an imputation of consensus for there to be genuine consensus: remember that there has to be the "objective possibility" that the expectations of the agent are treated as "valid" by the addressee, too.

There is a difference between merely *believing oneself to be* in a consensual relation with another being, and *actually being* in that kind of relation. It seems to be a necessary precondition for a relation to be consensual that the agents *believe* themselves to be in a consensus. But this is not *sufficient*. There are other criteria to meet. So it might appear that many of the cases in which people assume themselves to be in a consensus with non-linguistic beings are simply *mistaken*, and are not in an *actual* consensus, in spite of what they believe to be the case.¹³

We should not dismiss the possibility, however, that not all cases of imputed consensus with non-linguistic beings are of this erroneous kind. There is another reason to hold on to the idea of a non-discursive normative practice. From a genetic

¹³ As an example for this, many dog owners who have normative expectations with respect to their pet's behavior actually tend to mistake their dog for a linguistic practitioner, speaking to their dog as if it were a human being.

perspective, it does not seem easy to see how beings that cannot, in some rudimentary sense, *count on each other*, should ever come to use linguistic means to communicate. These considerations show that Habermas' link between consensus and language is by no means a matter of course. At the same time, it is clear that the idea of a pre-linguistic consensus is not without problems. There are problems on three levels: the epistemological, the moral or practical, and the ontological level. I shall briefly address each of these levels in turn.

From an epistemological point of view, it seems clear that any *verification* that there is an actual consensus presupposes language. Only linguistic practitioners can *confirm* that they are in a consensus. But this epistemological problem does not *per se* disconfirm the possibility of pre-linguistic consensus insofar as it is not inconsistent to assume an entity whose existence cannot be verified.

The second level is more difficult to address. An observation concerning the received literature on the sources of social normativity illustrates the point. There is much to learn from this literature about what entitles agents to have normative expectations that are addressed at the participants of linguistic practices, ranging from Christine Korsgaard's theory of the *Sources of Normativity* to Habermas' Discourse Ethics. In contrast to this, there does not seem to be *one single attempt* even to try to show why having normative expectations concerning the behavior of beings that are not capable of language might be anything else than totally unjustified. And there seems to be good reasons for this linguistic bias in the theory of social normativity. It is obviously rather problematic, if not outright absurd, to have such expectations. Why should somebody be considered to be *obliged* to a certain kind of behavior, if he or she has not committed herself to that norm? And how should such a commitment come about if not in some kind of discursive procedure?

I think that this objection is correct, but only as far as *morally laden* cases of normative expectations are concerned. For a consensus to be morally binding, there has to be consent, and consent implies endorsing a proposition, which is a linguistic matter. But not all normative expectations are morally laden. A dog owner who has normative expectations concerning his dog's behavior does not have to take his dog to behave *immorally* if it leaves its pile on the floor. There are "oughts" of weaker kinds. Let's not enter into a discussion about moral obligations here, and stick to the simpler question of whether there might be reasons to expect other agents' behavior normatively, if these agents are not linguistic practitioners and can therefore not be seen as morally obliged by consent.

Even if we accept these replies to the epistemological and the moral-practical objection, there are huge difficulties left. Many current philosophers claim that it is simply a *mistake* to apply the terms *action* and *practice* to pre-discursive behavior. This suspicion is nourished by current attempts to allow for the possibility of non-discursive normative practices. The most prominent example is Robert B. Brandom's theory. Brandom's explicitly stated aim is to maintain the conceptual possibility of non-discursive normative practices. This, of course, gives rise to the question: whence do those norms come from? Brandom's reply seems to be that they are simply instituted by *sanctioning behavior*. And that, Brandom claims, does not presuppose language. One does not have to use language to teach other people

manners. If necessary, the appropriate use of a stick will do just fine.¹⁴ The most obvious problem of this and other *sanctioning theories of social normativity* is that, for a behavior to be an act of *sanctioning*, the respective norm has already to be in place. A complex of behavior can be sanctioned only if it is measured against a norm. This is what distinguishes acts of sanctioning from mere aggressive behavior. Normative expectations are more fundamental than sanctioning behavior, because the latter presupposes the former. Thus again, there seems to be no place for normative expectations among non-linguistic beings.

Should the idea of pre-linguistic consensus therefore be dropped? I think it should not, and I believe that a closer examination of Weber's problem can show us why. If pre-linguistic consensus is possible, it has to be shown how Weber's problem can be solved without turning either to systems theory, or to a linguistic foundation for social theory. In other words, it has to be shown how consensual action can fit into Weber's framework of goal-oriented action. This is precisely what I will try to do in the remainder of this chapter.

A *caveat* is in order. There are two lines of objection against the following attempt. First, it can be argued that any kind of goal-orientedness presupposes the use of speech, and thus social normativity. Second, it can be argued that an instrumentalist conception of practical reason is *per se* anti-normativist, and cannot accommodate social normativity. Even though the first objection is more fundamental, I will only address the second in the following. In other words, I will simply take the possibility of pre-linguistic instrumental action for granted, and ask the question: how can the concept of consent be accommodated within a theory of instrumental action that does not count the normativity which is constitutive of consensual action among its preconditions? I gladly admit that, for complex calculations of means within multi-layered plans for action, this assumption is highly improbable. For more basic forms of goal-orientedness, however, this assumption might be granted the benefit of the doubt.

§44 Consensus and Commitment

From the above considerations emerges the following picture. The critical discussion of Luhmann's and Habermas' attempts at solving Weber's problem has shown that it is perhaps not entirely implausible to hold on to the following two basic Weberian intuitions: first, we have to give an *action-theoretic* account of consent, i.e. the structure of consent has to be accounted for *from the agents' points of view*, with an eye on the *reasons* those agents might have for whatever expectations they have. Second, consent is more basic than any form of linguistic procedure. Consent is not a genuinely linguistic matter, but rather among the presuppositions of linguistic communication.

¹⁴ Brandom (1997: 201): "I am indeed committed to the possibility of norms implicit in prelinguistic [...] practices. [...] The picture is that what proto-hominids could do before they could talk is to take or treat each other's performances as correct or incorrect by practically sanctioning them."

It is not necessary to assume that Weber's theory of consensus is superior to the theories of his critics in order to think that it might be useful to reconsider Weber's within the context of his own work while trying to find a solution to his problem. The question on which a second look at Weber's conception has to focus is the following: *what precisely is it that makes goal-oriented action such an implausible paradigm for a theory of social normativity?* Put in somewhat simplistic terms, the problem seems to be the following. If the agent's behavior is interpreted as instrumental action, and if it is believed that this is the fundamental way in which agents interpret each other's behavior, it seems entirely dubious why there should be anything like normative expectations.

Let's have a closer look at the matter. The problem is not that there is no place for normativity in instrumentalist accounts of practical reasons. Those accounts do imply normativity, but the kind of normativity at stake here is exclusively a matter of the relation between what an agent *wants* and what he or she actually *does*. If agent A wants x, he *should* (by virtue of the normativity of rationality) do y, i.e. choose the appropriate means to realize x (if there are no conflicting meta-preferences, or overriding desires). The normativity of y is *hypothetical*: it is something A should do *insofar* – and *only* insofar – as he or she wants x (or anything else that is best realized by y). Within this setting, there is no telling *what it is* that A should *want*. It is true that Weber is not Hume, for whom “reason is (...) the slave of passion”. In contrast to Hume's passions, Weber's goals are subject to rational constraints – but only insofar as there are inconsistencies concerning the values of the agent. If consistency of the entire range of consequences with the agent's values is integrated into the concept of goal-effectiveness, Weber's famous claim that there is no rational evaluation of one's values can be transferred to one's goals (i.e. insofar as goals are value-consistent effects of the behavior that are weighed against any unintended consequences).

The decisive point, however, is yet another one. The range of purely instrumental normativity is limited to a purely monological space: instrumental rationality is a matter of the solitary relation between an agent and his or her goals. It ties the agent's behavior to his or her ends, and not one agent's behavior to another agent's. There is, it seems, nothing social about this normativity. Thus this kind of normativity seems utterly unfit to explain how rational agents should come to have normative expectations concerning each other's behavior. Instrumental agents might think of each other as rational or irrational, depending on the effectiveness of their pursuit of their goals. Insofar as there is no reason to think that other agents *should* pursue whatever goal they happen to have, there seems no reason normatively to expect them to pursue their goals *rationally*, either. It seems that the only person of whom a purely instrumentally rational agent can expect rational behavior in a *normative* way is the agent him- or herself, because rational agents cannot be indifferent as to the question of whether or not *their own goals* are achieved. Thus instrumentally rational agents cannot take a purely cognitive stance towards their own (future) behavior. When their mind is set on a goal, they have to *expect* themselves to pursue their goals effectively. It is true that agents might *take into account* the weakness of their own will, and might suspect that they will not be able to follow through with their own

plan. Thus there are cognitive components in their expectations concerning their future behavior. But their expectations cannot be *entirely* and *exclusively* cognitive. There *has* to be a normative component, at least insofar as the agent's goals are not just desired states of affairs, but *conditions of satisfaction of the agent's intentions*, i.e. whatever has to be the case for an agent to have done what he or she intended to do. In other words, to have an *intention* implies to have *some* normative expectations towards one's own behavior.

It appears, however, that within an instrumentalist account of practical reason the agent's own behavior is the *only* suitable address for normative expectations. The only genuinely normative relation an instrumentally rational agent has is the solitary relation to him- or herself.

This does not, of course, mean that instrumentally rational agents will be entirely *indifferent* to other agents' irrationalities. Where this seems useful in the light of their own plans, they will encourage other agents to be rational, too. But just as other people's values have to be taken into account, so do other people's irrationalities. There is no reason to *count on* each other's rationality, and to act on normative expectations regarding each other's rational behavior.

These considerations lead to a view of Weber's problem that is very different from either Parsons' or Habermas'. The main problem of the attempt to integrate a theory of social normativity into an instrumental account of practical reason is not the focus on action (as systems theory suggests), nor is it a matter of the focus on instrumental success as such (as Habermas believes). Rather, the problem can be narrowed down much further. It lies in the way in which, within Weber's account at least, action (in terms of meaningful behavior) is related to instrumental success (in terms of conditions of satisfaction of intentions). Just as Weber conceives of action as a matter of solitary agents, he thinks that the goals in question have to be solitary agent's goals, too. This *individualism about goals* is particularly obvious in the first introduction of the concept of instrumental rationality in the second paragraph of *Economy and Society*. Instrumentally rational action is "determined by expectations as to the behavior of objects in the environment and of other human beings; these expectations are used as 'conditions' or 'means' for the attainment of the actor's *own* rationally pursued and calculated ends" (WG 12). The emphasis is mine, for my concluding question is the following: could it be that *this* is the cause of Weber's problems: neither his focus on action, nor his alleged instrumentalism, but his *individualism about goals*?

Now a days, the place where the role of goals in action is most hotly debated is Rational Choice Theory. In this context, Amartya Sen has repeatedly questioned a more or less tacit assumption of the orthodox model of practical reason that has been accepted at least from Max Weber's days up to current Rational Choice Theory. Sen calls the assumption in question *self-goal choice*. Self-goal choice does not imply that the agent's goal is focused on his or her own self-interest. The assumption that agents aim at furthering other people's well-being is compatible with self-goal choice. The limitation at stake here is of a different nature. Self-goal choice is the claim that agents have to be interpreted as pursuing *their own* goals. But how could

this ever be otherwise? How could agents choose anything else than their own goals? How could goals be anything other than the agent's own?

Sen claims that there is a type of action, which he calls *committed action*, which is in a sense *goal oriented*, but not towards the agent's own goals. Sen is clearly aware of how strange and indeed counterintuitive his claim might appear. For even in paradigmatic cases of altruistic action, self-goal choice does not appear to be violated, because the altruists' own goal seems to be to benefit other agents, which is his own (altruistic) goal. As Sen himself puts this objection to his claim, "it might appear that if I were to pursue anything other than what I see as my own goals, then I am suffering from an illusion; these other things *are* my goals, contrary to what I might believe." (Sen 2002: 212)

Thus it is hardly surprising that most critics do not accept Sen's claim.¹⁵ I think this defense of Rational Choice has a point against Sen, but only as far as the only alternative to the agent's individual goals are other agent's individual goals. In contrast to this, it seems to me that the merit of Sen's claim is to point our attention to another class of goals. The fact that agents cannot pursue other people's goals without making them their own does not mean that all goals which agents pursue are their own goals. Agents do not just have their own goals. They have *shared* goals, too.¹⁶ Interestingly, it seems that the early Parsons had similar ideas. In his critique of the "atomism" of both utilitarian and contractual versions of practical reason, Parsons has made the very same point, emphasizing "that men's ends should not be separate, and either forcibly restrained or miraculously compatible, but in fact, in a given society, *held in common*." (Parsons 1934: 158; my emphasis)

It seems that the existence of shared goals opens up an alternative perspective on the relation between instrumental action and social normativity, one that makes a solution to Weber's problem of having to base a theory of social normativity on an instrumentalist theory of action possible. Where agents pursue a shared goal, instrumental normativity breaks out of the cage of the solitary relation between an individual agent and his or her own goals, and takes on a social meaning. As far as agents pursue shared goals, the instrumental rationality of other agents concerns them just as directly as their own instrumental rationality does. Just as solitary rational agents cannot be indifferent as to the achievement of their own goals, and cannot take a purely cognitive stance towards their own future behavior, they are now in normative relations to each other. If A's individual goal x_i is a *reason for him* to do y (as far as y is the best means to realize x_i), agents acting in pursuit of the *shared goal* x_s can mutually expect rational behavior (i.e. y) *from each other*, because y is the best means to realize x_s .

In other words, shared goals appear to be the point at which merely instrumental requirements turn into social normativity. Just as the relation of individual agents to themselves cannot be limited to cognitive expectations, the relations *between* agents acting in pursuit of shared goals have to include a normative element, too. They have to expect each other to choose the suitable means *normatively*. They do

¹⁵ Cf. e.g. Pettit (2002).

¹⁶ For a similar reconstruction of Sen's concept of committed action see Anderson (2001).

not just take each other's behavior into account; in the way the archer takes into account the wind when she aims at her target.¹⁷ The behavior of the other agents is not just another *restriction* that has to be taken into account within one's own calculation of suitable means. And much less do agents who act in pursuit of a shared goal simply manipulate each other as means to their own goals, as Jean-Paul Sartre thinks. Rather, these agents *count on each other (as well as on their own future selves)*, and their shared goals provide them with a point from which they can critically assess other people's behavior as well as their own.

If pre-discursive instrumental action is possible: why should this be limited to the pursuit of *individual* goals? Why should pre-discursive instrumental action not extend to the pursuit of shared goals? Especially in light of recent findings in child psychology, which has shown how important shared intentionality is even for infants, which are not yet capable of speech, and how deeply rooted the motivation for cooperation is in our psychologies (cf. Tomasello 1998), this does not seem implausible at all. If this is right, then the solution of Weber's problem is to be found in a theory of shared goals. And if the primary role goals have for action is that of conditions for the satisfaction of intentions, we need an account of *shared intentions* to solve Weber's problem. The most important amendment to be made to Weber's action theory is this. If it is a characteristic of instrumental action that it is "determined by [cognitive] expectations as to the behavior of objects in the environment and of other human beings" which are "used as 'conditions' or 'means' for the attainment of the actor's *own* rationally pursued and calculated ends" (WG 12), consensual action is, at the most basic level of the phenomenon, determined by *normative* expectations regarding the contributions of other agents for the agents' rationally pursued and calculated *shared* ends.

¹⁷ Oddly enough, this is what Pettit seems to think; cf. Pettit (2005: 20).