

Raimo Tuomela/Kaarlo Miller

We-Intentions and Social Action

Abstract: In the paper "We-intentions and Social Action" conceptual issues related to intentional social action are studied. By social actions we here mean actions that are performed together by two or more agents. The central concept of we-intention is introduced and applied to the analysis of simple social practical reasoning. An individualistic analysis of the notion of we-intention is proposed on the basis of the agents' I-intentions and beliefs. The need and indispensability of we-intentions and we-attitudes in general in a theory of intentional social action is emphasized along with the fact that we-intending leads to action in suitable circumstances.

1. Introducing Social Actions and We-Intentions

In philosophical action theories surprisingly little attention has been devoted so far to actions jointly performed by several agents. Besides game theory, there are few systematic and detailed accounts of the strategic interaction of agents, let alone broader and more realistic philosophical theories of social action. An account of joint action can be considered to be a part of general action theory, and it also creates interesting conceptual and theoretical problems of its own. In this paper some of these specific problems will be discussed, and a sketch of some of the essentials of a systematic theory will be presented (see Tuomela 1984 for a full presentation).

The theory of social action to be discussed is an extension of the so-called purposive causal theory of single-agent action (developed in Tuomela 1977) to the multi-agent case. A central feature of this theory is the causally effective role of the agent's active and executive intendings (or willings) in producing behavior. Our main concern in this article is to explicate the central concept of we-intention, which intentionally performed joint actions conceptually involve.

Most of our actions are social in the wide sense that they conceptually presuppose the existence of other agents and various social institutions. Of actions that are social in this sense, some are performed by single agents while the rest are either performed jointly by several agents or performed by collectives of agents. An action's being performed jointly by several agents and its being performed by a collective of agents are not the same thing. Actions that are performed jointly by several agents we shall call multi-agent actions and also social actions proper. Multi-agent actions are of course social in the above wide sense. Examples of such multi-agent actions would be two or more agents' (jointly, rather than separately) carrying a table upstairs, playing tennis, toasting, conversing, etc.

We argue that a joint action, of type X , say, performed by some agents, say A_1, \dots, A_m , involves that each of these agents does something - his part, as we may call it - when the agents jointly perform X . If we let X_i , $i=1, \dots, m$, stand for A_i 's part action or component action, we may ask how the performances of the X_i 's should be put together, as it were, so that we get a performance of X . This is a non-trivial problem as obviously the agents could, for instance, each of them build a house or sing a song without it being the case that they jointly build a house or sing a song. So what kind of interaction or relatedness should we require here?

A central claim here will be that all intentional joint many-agent actions will involve some relevant we-attitudes, viz. we-intentions and mutual beliefs plus the we-proattitudes underlying them. (For weaker notions of social action, e.g. actions making essential reference to some other agents, see Tuomela 1984, Ch. 5.). In other words, we claim that the 'sociality' or 'social relatedness' central to people's acting together in a central sense comes from or even consists in their relevant we-attitudes. This is a point of view from which one can approach such problems as the typology of multi-person actions, social practical reasoning, social control, and most importantly communicative action. Communication in its fullest and most interesting sense may be taken to involve so-called reflexive intentions and Gricean mechanisms.

As our aim in this paper is to concentrate on the concept of we-intention, we will for the most part ignore the overt aspects of the 'jointness' of social action. The issues here bypassed include e.g. the important problems of 'putting together' the agents' bodily behaviors and the results and consequences of the X_i 's as well as the various needed individual (causal or conceptual) generational relations. A reader interested in structural analysis of social action is referred to Tuomela (1984), where the structure of social action and the generational relations are studied in detail.

In the case of actions that the agents jointly intentionally perform it seems necessary that they share a common intended goal, normally (but not necessarily) the goal to perform the total action, say X. Indeed, they must share a relevant group-intention, viz. a we-intention expressible by "We shall do X" (or "We will do X"), even if this intention need not be formed prior to action.

We may now ask why a common intended goal or a group-intention should be present in intentional joint action. As our intuitions about social actions are not very sharp, some amount of stipulation in drawing distinctions will have to be involved. But given this, what we claim is that at least a full blown notion of intentional social action should be taken to involve such a group-intention or we-intention on the part of every agent. This is of course very tentative; we will below present arguments and examples for the introduction of we-intentions. We will also argue that we-intentions are indispensable for social practical reasoning.

Let us concentrate on such a full blown concept. Suppose some agents $A_1, \dots, A_k, \dots, A_m$ (perhaps repeatedly) jointly perform X, e.g. sing a song or play a game of cards, and do it intentionally. We cannot say they did it fully intentionally, viz. intentionally as a collective (at least if the collective is antecedently unorganized), if any of them lacked a relevant group-intention (usually one to do X) expressing the agents' common goal, even if each agent would have performed his part intentionally. Had, e.g., A_k performed his part of X intentionally but without sharing the other agents' common goal he would not have intentionally acted jointly with them, we may say. Because of this the agents would not intentionally have acted as a group, and so their social action would not have been fully intentional.

Consider next an example where agent A gave lethal poison to C to kill him and where B did the same, too. Both A and B gave an amount sufficient to kill C, we assume. A and B acted without knowing of each other's intention. Can we say that they intentionally jointly killed C? Obviously not - rather we say that each killed C (or contributed to the killing of C) separately. Why is this so even if A and B in a sense both had the same intention to kill C by poisoning, and also acted on this intention? This is so essentially because they lacked the mutual awareness of each other's intentions. But when we add this requirement of their mutual awareness of each other's plans we in fact arrive at we-intentions almost in the sense they will be below analyzed by the schema (WI) (apart from the fact that in (WI) I-intentions are conditional) and at the requirement of acting on such we-intentions.

We may also view the situation from the point of view of a collective's action. Why does not a mere shared intention (rather than a we-intention)

to do X (or something else relevant) suffice? Suppose things go wrong when the collective starts doing X. For instance, one of the agents may fail to do his component action. Then, ideally at least, the others will help, exert pressure, and do whatever they think is necessary for the collective to do X. This again indicates that it must be believed by everyone in the collective that everyone else shares the relevant intention leading them to do X. To the extent these beliefs are hierarchically justifiable we may speak of the agents' mutual beliefs here. A mutual belief that everyone in the collective intends to do X ideally consists in everyone's believing that everyone intends to do X and everyone's believing that everyone believes that everyone intends to do X, and so on theoretically ad infinitum (even if in actual practice only two or three layers may be needed). The iterability of "everyone believes that" can be regarded as giving justification to the lower degree beliefs.

2. Definition of We-Intention

The concept formation programme followed by us is basically functionalist: social concepts are to be characterized in terms of their functions or roles in the theories where they are introduced and employed. This programme is explicated in detail in Tuomela's recent book (1984, Ch. 2). Here it suffices to state the following: The general concept formation programme is called Conceptual Individualism, according to which holistic social concepts are constructed out of individualistic concepts, especially action concepts. This presupposes the availability of relevant (individualistically analyzed) concepts of other-regarding attitudes (we-attitudes, especially we-intentions) in a viable theory of intentional social action. We-intendings are other-regarding intendings of an individual reflecting the concept of group ("us"). Thus our account in a sense uses holistic ideas (such as group-notions) but analyzes them (relatively) individualistically.

We can now try to summarize the central metaphysical content involved in the notion of we-intention. In our analysis it is individuals who intend and make inferences, but what is common is their goal; and it is just this goal that is reflected in their we-intentions. The agents' 'intendings' are not strictly speaking 'shareable'. What they share is their goal, and thereby they share a common content involved in their we-intentions. Nobody else can perform the agent's intending for him, like someone else can perform (share etc.) his work. The agents participate or collaborate in action, not in intending. Yet the intendings are about a joint end, directly or indirectly.

Given the motivation for the need of we-intentions presented in Section 1 and above, the concept of we-intention can be defined in terms of conditional I-intentions and I-beliefs as follows, where X represents an ante-

cedently characterized social action type (in the sense of (8) of Tuomela 1984, Ch. 5):

(WI) A member A_i of a collective G we-intends to do X if and only if

(i) A_i intends to do his part of X , given that he believes that every (full-fledged and adequately informed) member of G or at least that a sufficient number of them, as required for the performance of X , will (or at least probably will) do his part (their parts) of X .

(ii) A_i believes that every (full-fledged and adequately informed) member of G or at least that a sufficient number of them, as required for the performance of X , will (or at least probably will) do his part (their parts) of X .

(iii) there is a mutual belief in G to the effect that (i) and (ii).

Let us next explicate what is contained in the clauses (i)-(iii). In clause (i) A_i 's intention is supposed to be a plain intention (I-intention) to do his part of X .² His I-intention is conditional on his belief that the others in G will do their parts of X . Clause (ii) states that this condition is fulfilled. From clauses (i) and (ii) one can validly infer that A_i unconditionally intends to do his part of X . This expresses also a sound pattern of practical inference or intention deconditionalization for A himself, at least if *akrasia* is barred. And finally clause (iii) states that it is mutually believed in the group or collective G that (i) and (ii) and via deconditionalization that A unconditionally intends to do his part of X . The fact that the agents have the mutual belief makes their we-intentions intersubjective. (It is a somewhat tricky question how such mutual belief is to be best analyzed, for it is in many cases too strict to require that every member of G - rather than e.g. every typical operative member, for instance - must have such a belief, which, moreover, should be suitably iterable to yield higher order beliefs; see Tuomela 1984, Ch. 7 on this.) Thus, anyway, each A_i does not merely independently happen to share with the others an intention to do something; he is in addition aware of the others' intentions, and believes that the others are aware of his. (The notion of mutual belief is explicated in detail in Tuomela's mentioned book 1984, Ch. 7; also cf. Lewis 1969.)

Let us apply (WI) to a concrete example to show the role and need of the mutual belief clause (iii). Let X be the joint action of carrying a table upstairs by two persons. Let X_a be A 's part, and X_b B 's part. What X_a and X_b contain need not be specified in advance. For brevity B_a is used to denote A 's belief and I_a to denote A 's intention. Let us also adopt the convention that $B_a X_b$ and $I_a X_a$ mean that A believes and intends respectively 'in the present tense' that X_a and X_b be realized in the future, at the appropriate time.

From (WI)(i) and (WI)(ii) we get:

- (a) A intends to do his part ($I_a X_a$).
- (b) A believes that B will do his (viz. B's) part ($B_a X_b$).

It is also assumed that:

- (c) A believes that he cannot do his part (X_a) by himself.
- (c*) B believes that he cannot do his part (X_b) by himself.
- (d) A believes that he cannot do the total social action (X) by himself.

A can of course intend to do the carrying all by himself (if he considers that possible), but this is a different action (and respective intention) from both his doing his part (X_a) and from the joint action (X). In (c) and (c*) it is also taken for granted that the agents believe they cannot do their respective parts alone. It might happen that A believes that he can do alone everything that amounts to X_a . This, however, is not the same action (nor same respective intention) as his doing X_a as his part of X (cf. also notes 1) and 2) below). If we concentrate on A's beliefs only, from (WI)(iii) we get:

- (e) A believes that B believes that he (A) will do his part ($B_a B_b X_a$). Or at least $\neg B_a \neg B_b X_a$.

Now we are ready to ask if the truth of (e) is necessary for 'A to we-intend to do X' (the analysandum of (WI)). If we sum up (a), (b), and (e) using A's positive second order belief of (e), we get the following 'state-description':

- (f) $I_a X_a \wedge B_a X_b \wedge B_a B_b X_a$

To answer our question let us first assume that A lacks the second-order belief of (e). - Are the following cases possible (consistent)?

- (g) $I_a X_a \wedge B_a X_b \wedge \neg B_a B_b X_a$

- (h) $I_a X_a \wedge B_a X_b \wedge B_a \neg B_b X_a$

- (i) $I_a X_a \wedge B_a X_b \wedge B_a B_b \neg X_a$

In (g) the stronger version of (e) is not satisfied, and in (h) and (i) not even the weaker version of (e) is satisfied. ($\neg B_a \neg B_b X_a$ and $\neg B_a B_b X_a$ can be valid at the same time, but this is not the case with $\neg B_a \neg B_b X_a$ and $B_a \neg B_b X_a$ nor with $\neg B_a \neg B_b X_a$ and $B_a B_b \neg X_a$.) Let us consider the case (i) as an example.

In (i) it is asserted that A believes both that B will do his (B's) part and that B believes that he (A) will not do his part. In our example this means that A believes both that B will participate in the carrying and that B believes that A will not participate. We can transform (i) to the following:

$$(i') \quad I_a X_a \wedge B_a (X_b \wedge B_b \neg X_a)$$

Some weak rationality assumptions may be required in the transfer from (i) to i'), but we can ignore them for our purposes. Is (i') possible or at least 'doxastically defensible'?

Let us begin by considering the content of A's belief in (i'). If A believes that B will participate, he of course believes that B can participate (A cannot believe what he considers impossible). On the other hand $B_b \neg X_a$ entails (by virtue of (c*)) that B believes he cannot and thus will not participate. Thus A would believe both that B will participate and that B believes B will not participate: $B_a (X_b \wedge B_b \neg X_b)$. Provided that A does not give up his belief, $B_a B_b \neg X_b$ entails that A believes that B will not (intentionally) participate! (He can of course believe that B will do X_b by mistake, but this case does not concern us here. Thus $B_a (X_b \wedge B_b \neg X_a)$ in (i') leads to the inconsistency $B_a (X_b \wedge \neg X_b)$.)

A's belief that B will not participate, besides contradicting (WI)(ii) and (b), implies that he does not have the unconditional intention ($I_a X_a$ of (i')), and this in turn that he does not we-intend to do X (in the circumstances and with B).

We have considered the case (i) and shown that the assumption $B_a B_b \neg X_a$ means that (i) and (ii) of (WI) will not be satisfied. This indicates that at least some degree of mutual awareness is necessary. Thus we must assume that A's belief that $B_b \neg X_a$ must be absent.

In the above example the relevant beliefs were de re. Concerning A_i's de dicto belief the above result could be generalized as follows: If A_i were to believe that every other possible member in G believes that A_i will not do his necessary part of X, this belief of A_i's would be incompatible with (WI)(i), (WI)(ii), and his we-intention. In this case we must assume that A_i believes that his action is considered a necessary part of X.

On mere conceptual grounds one cannot say much relevant of the above examples. One can, however, look for broader factual support for the mutual belief requirement in (WI). In many cases group pressure is directed towards A_i in order for the collective to be able to do X. For that pressure to be effective A_i has to be aware of the relevant members' intentions and they must be aware of his. The purpose of the group

pressure is to 'reinforce' the relevant we-intention(s), but it also presupposes some mutual awareness of the intended goal. Group cohesion, altruistic behavior, e.g. helping, often also presuppose strong mutual beliefs (cf. schema (4) of section 3).

There is still one central point to be noted in connection with (WI). The clauses (i)-(iii) can be used to show that A_i unconditionally intends to do his part of X at least partly because he believes that everyone in G, i.e., each of us will do (his part of) X. This is central for the justification of the intention. (We can of course think of anomalous cases, where A_i intends to do X (or his part of X), not because, but despite the fact that he believes that the others will do X, because he cannot but collaborate. And yet his we-intention would satisfy the definiens of the schema (WI).)

3. We-Intention and Practical Reasoning

In many cases of intentional joint action each A_i in G can be assumed to reason according to the following simplified schema, where (i) expresses a we-intention and (ii) a belief of A_i :

- (1) (i) We will do X.
- (ii) I am one of us (viz. the collective = $\{A_1, \dots, A_m\}$).
- (iii) I will do my part of X.

We can conclude that A_i 's we-intending to do X commits him to do anything in his power he considers necessary for him to do in order for G to do X, at least if "us" in (ii) refers to the operative members of G. So if X_i is all that A_i considers necessary for himself to do (his 'duty', we may say), we get the following idealized schema (here expressed for the first person case):

- (2) (i) We will do X.
- (ii) Unless I do X_i , we cannot do X.
- (iii) I will do X_i .

(2) is a practical inference schema one can refer to when justifying A_i 's intention to do X_i . When (2) is a practical inference schema with true premises and a true conclusion, premise (ii) expresses what A_i regards as his (minimum) action part in X. (It should be noted that (ii) need not be true in the case of all social actions.) His we-intention commits him to do at least X_i . As the definiens of (WI) can be taken to satisfy the inference schemas (1) and (2), it is easy to see that (WI) could also be applied to a naturalistic analysis of some moral concepts (e.g. duties) in the sense that one's duties can be relativized to (historical and contingent) social group purposes.

Thus, when reasoning according to (2), each A_i intends to do everything he believes necessary for him to do (i.e. X_i) for the total action X . If normal conditions for each intention-execution obtain, and if all the A_i 's are right in their beliefs ((2)(ii)), their intendings to do the X_i 's will result in their jointly doing X .

It is of course not clear offhand what the normal conditions are supposed to include. In the case of single-agent action one position would be to take the successful performance of the intended action itself as the requirement and criterion of the obtaining of the normal conditions. Provided that the agent does not change his mind, this view entails that if the agent failed to perform the intended action, this indicates that some normal condition(s) did not obtain, i.e., the cause of his failure is to be found in the non-obtaining of some factor in the class of normal conditions. According to this view having an intention implies the intended action within intended time in normal conditions (e.g. Sellars 1980 adopts this kind of view). If we understand normal conditions in this strong sense, then our above thesis can be accepted as true.³

In normal conditions A_i 's intention to do X_i results in his doing X_i . (In this respect intentions differ from mere wants or wishes. We cannot, however, here explicitly argue for this principle (cf. e.g. Tuomela 1977, Ch. 7; Sellars 1980).) When each A_i has performed his X_i , and if each A_i was right in his belief that his X_i contains everything necessary for him to do for X , it follows in normal conditions that they have done X (or will succeed in doing it). As the conjunction of all the necessary conditions of X is (tautologically) a sufficient condition of X , we can say that the conjunction of each necessary single-agent component X_i is sufficient for X as far as A_i 's actions are concerned.

It is of course possible that, although all the necessary X_i 's are realized, X is not realized. If this would turn out to be the case, we can say that X was not realized because the normal conditions were not satisfied.⁴ The A_i 's have done all that is required of them to do X . If there would still be something left for the A_i 's to do for X , this would contradict the assumption of the correctness of belief, i.e., there must have been at least one A_i whose belief of his necessary action component for X was mistaken.

When the performance of joint action is concerned, the normal conditions have to obtain on two levels, viz. on the level of making individual performance possible and on the level of making joint performance possible. These conditions could also be called (internal or external) individual action opportunities and joint action opportunities, respectively. They do not always coincide. One of the opportunities of the latter (but not the former) kind for each A_i to participate in a joint action X is that every

member of G (or at least a sufficient number of them) will participate in X. When we-intending to do X, each A_i believes that there are opportunities (or an opportunity) on both levels, or, to be more accurate, it is a conceptual truth that each A_i does not believe (is not confident) that there is not an opportunity for either his individual performance or for their joint performance.

It is also important to note that according to (WI) A_i believes that the others will (or probably will) their parts of X. His believing that the others merely intend to do their parts of X is not enough. To be sure, normally A_i believes that the others also intend to do their parts of X. Yet the possibility that A_i believes that the others will do their parts unintentionally (e.g. by mistake) is not excluded in (WI).

Consider the following example as an argument for the requirement that A_i cannot merely believe that the others intend to do their parts in the case of intentional acting. Suppose that A intends to marry B, given that C intends to marry B, and C intends to marry B, given that A intends to marry B, and that this is mutually believed. Neither A nor B can be said to have a corresponding we-intention (in a monogamous society), and yet this is a completely possible situation. So it is also possible for both A and C to intend unconditionally without believing that the other intends. Here we have a case where there is an individual action opportunity (and an I-intention opportunity), but no joint action opportunity (nor a we-intention opportunity). The conventions of the society define what counts as marriage (and marriage ceremonies) and therefore, although there are physical individual and joint action opportunities, there is no conventional joint action opportunity for A and C. This example illustrates also the fact that having a common intended goal, although necessary, is not sufficient for having a corresponding we-intention. More generally, we can say that if an agent thinks the others intend to perform their parts of X but yet is convinced that they will not be able to realize their intentions, he will not (unconditionally) intend to do his part (which we took to be inferable from (WI)).

According to the premise (ii) of (2), the agent considers his action X necessary for the total action X, i.e. in his opinion his doing X is a necessary condition for their doing X. It follows that the result of his action (i.e. what he brings about) is a necessary condition of their joint goal when (ii) is true. This necessary conditionship relation is neither objective nor absolute, but 'practical'; it is supposed to hold true (a) in the agent's opinion, and (b) in the circumstances. In our above discussion of the schema (2), each X_i was thought to comprise everything each A_i believes necessary for himself to do for X. His belief was also supposed to be true. And we saw that the truth of these presuppositions entails that X will be realized in normal conditions.

If A_i does not change his mind about his we-intention (premise (i)), he has to do something, in any case there is something which is practically necessary for him to do for X . (Thus cases where some other agents represent the original agents in X are not considered here.) In our discussion above, X_i was supposed to contain everything A_i regards as necessary for himself to do for X . Assuming that A_i does not change his mind about premise (i), we are still left with two possibilities in connection with (2): A_i may not after all accept its premise (ii) (and its conclusion) and thus he may either intend to do less than X_i (i.e. less than his original share of X), or he may intend to do more than X_i . (He may of course intend to perform a totally different subaction X_j , but this case does not concern us here.)

If A_i intends to do less than X_i , and if in the other agents opinion he has to do the whole original X_i , the other agents think - contrary to him - that this does not 'relieve' him from his practical necessity to do that part of X_i (a part of a necessary condition is itself a necessary condition), but in addition to this he has to see to it that the 'residue' of X_i will be taken care of. This is a typical case of social action, and it usually means negotiations, and, if A_i is successful in his claim, a new 'division of labor' will result, in which at least one other agent has his practical necessities increased. This new situation (the newly acquired 'balance' of intentions and goals) is normally in other respects equivalent to the original situation.

Let us next consider the second case, where A_i intends to do more than this X_i . Unlike the first case, no change in A_i 's practical necessities for the achievement of their common goal is involved here. A_i 's we-intention and his belief still commit him to do X_i , but in his opinion it is not necessary for him to perform the 'surplus action(s)' for their doing X . This kind of example is also a typical case of social action. When acting together the agents often do more than they consider necessary to do for their joint action. For example many games and other informal leisure activities are like this; also cf. morally supererogatory actions. But also in these cases it is normally necessary that each agent does something, his necessary 'core action' when he we-intends to do X (i.e. intends to do X together with others) and carries out that intention. In cases like this A_i 's 'surplus action' is not essential to their achieving their goal, and in explaining this action one often has to look for some other intentions or motives than A_i 's mere we-intention to achieve their common goal.

Finding true premises for a schema like (2) so that each A has true beliefs and performs his X_i , and no more (nor less) for their total action X ; means finding an application of a principle of economy to the action X and to the collective in question. It is also compatible e.g. with Max Weber's characterization of the ideal type of rational bureaucratic action.

Schemas like (2), when expressed in the third person case, can be used to explain the agent's action. His we-intention, together with his beliefs about the situation explains why he performed his action. It is worth noting here that A_i 's we-intention to do X together with his belief that X_i is sufficient (given that the others will perform their share) for their doing X does not entail that A_i intends to do X_i nor that he actually does X_i in normal conditions. He can very well choose some other action X_j which he may also consider sufficient. In some situations A_i may even consider that his practical necessity for their doing X consists of a disjunctive action, i.e., he believes that he has to perform one or another of the disjuncts for their doing X. If one wants to explain why A_i chose just this disjunct, citing his we-intention and belief is no longer sufficient, and one has to look for other (true) intentions or motivational factors to explain this fact.

The schema (2) is an idealized pattern, and usually intergroup communication and various adjustments are required, before the agents reach agreement on their component actions. It shows, however, how a traditional practical syllogism pattern can be applied to the analysis of joint action. The premises (i)-(ii) entail that X_i is practically necessary for A_i , he has to do at least X_i (or else he would believe that they cannot do X which contradicts his we-intention expressed in premise (9)). And we have shown that if normal conditions obtain, and if each A_i is right in his beliefs (premise (i)), X will result, too.

When reasoning according to (2) the agent infers to or derives his I-intention from his we-intention, and carrying out the former is (in the agent's opinion) a necessary means for carrying out the latter. Here we can make a distinction between direct and indirect I-intentions. The former are his plain or 'egocentric' intentions, the latter are his we-derivative I-intentions (cf. Sellars 1980, 99). This distinction, however, is not absolute, and our next practical syllogism schema shows how also A_i 's we-intention can be I-derivative.

An inference schema partly analogous to (2) can be used to illustrate how a we-intention may be acquired. This can be done by simply changing the places of the intention to do X and the intention to do X_i in the premises of the schema (2). The crucial question is now what these new premises will yield as their conclusion. We propose the following schema:

- (3) (i) I will do X_i .
- (ii) Unless we do X, I cannot do X_i .
- (iii) I will do whatever I consider necessary to ensure that the others in G will do their parts of X.

From (3) one can see that A_i 's intention to do X_i (X_i as his 'desire', we may say), and his relevant belief commit him to ensure that the others will participate in X . When this task is accomplished, i.e., when A_i as a consequence has acquired the belief that each member of G will do his part of X , we can almost say that he has acquired a we-intention to do X , for he can be considered to come to satisfy clauses (i) and (ii) of (WI), interpreting (iii) of (3) in the strong success-sense allowing A_i to come to believe that the others indeed will do their parts. If in addition the successful carrying out of (iii) of (3) gives the required mutual beliefs, our (WI) is satisfied. Thus we may also proceed the other way and instead of considering what an already present we-intention contains we rather study the case where A_i 's plain or 'egocentric' intention (or even desire), together with his belief about the situation, requires him to do anything he considers necessary to ensure that each member of G will do his part of X . In this case A will do anything he regards as necessary to make sure, and thus to acquire the belief, that the others will participate in X , and if he is fully successful (in the sense of making also the others to mutually believe what he here intends and believes), he acquires the we-intention to do X .

For an example, think of a case where A_i intends to take a taxi to attend a seminar outside town, but he does not have enough money to do it by himself (he does not like walking and no other transportation is available). He considers that unless he travels together with his colleagues, he cannot go by taxi. Therefore he has (if he does not give up his original intention to take a taxi) to see to it that the others will share the trip with him. Thus he will then do whatever he considers necessary for that end, e.g. to telephone his colleagues and ask them to participate in the trip. As a result he may acquire the belief that the others will participate, and thereby he has acquired the we-intention to take a taxi. We-intentions often originate from the agents' personal wants or desires. The agents' duties, on the other hand, often have their roots in social group-purposes, and thereby in we-intentions.

Not only does the we-concept figure in simple reasonings of the above kind but also in more complex ones. Thus we may consider the following practical inference schema, containing intention- and belief-expressions, to represent my reasoning in many cases of social action. Here my being one of us serves as a partial, other-regarding reason for my doing one or more other actions, such as Y below, less directly connected to our jointly doing X . So I accept:

- (4) (i) We will do X .
- (ii) A is one of us.
- (iii) I intend (to bring it about) that A will do whatever I regard as necessary for him to do in order for us to do X .

- (iv) I will do whatever I regard as necessary to bring it about that A does whatever I regard as necessary for him to do in order for us to do X.
- (v) My doing Y (e.g., teaching A to do something) is such a thing.
- (vi) I will do Y.

We have above given reasons for the need to apply a concept like we-intention in a theory of social action. Next we have to comment briefly on how the I-intentions and we-intentions can be related (as to their contents) in the total social action X.

(WI) is assumed to concern a we-intention to perform a social action X. A social action type X can, relative to a context, technically be analyzed as a conjunction $X_1 \& \dots \& X_m$, where X_i represents agent A_i 's part of the social action X. (How X comes to be identical with $X_1 \& \dots \& X_m$ may in some cases be due to the constitutive meaning postulates, or the like, of the language or it may be due to, e.g., highly contextual features. Analogous remarks apply to the problem of how A_i will be assigned just X_i .) The definiens of (WI) entails that each member of G intends to do his part of X. If, for instance, X is playing tennis, cleaning the house, picking berries or conserving energy then the X_i 's coincide and we may use the same verbal phrase, e.g., 'conserving energy' in our example, both for X and its parts, the X_i 's. In this sense each member of G may be said to both we-intend and I-intend to do X. If, on the other hand, X is, say, an irreducible multi-agent action (such as getting married) so that each agent A_i is assigned a component action X_i , perhaps even so that $X_i = X_j$ for some j, then it seems problematic to say that an agent intends to do X, as no agent A_i alone can do X. Therefore the definiens of (WI) says that A_i intends to do his part of X rather than X itself.

We have above emphasized that A_i 's we-intention normally leads to action.⁵ This feature could attract the attention of social psychologists studying the action of groups. Social psychologists in general, when discussing group purposes, have not clearly differentiated between intentions, wants, other motivational factors and externally given goals, although all these have their own distinctive conceptual features. (For a relatively detailed analysis of group goals cf. Shaw 1981, Ch. 10 and Cartwright/Zander 1968.)

The we-attitudes we have presented can be compared with G.H. Mead's famous distinction between the passive 'me' (that represents other people's ideas and social role expectations of the individual), and the active 'I' (that represents the impulsive aims and spontaneity of the individual). The dialectics between these two constitute the individual's real 'self'. This distinction could be complemented with a we-concept; the active 'we' that

applies to the individual as a group member; and 'us', its passive counterpart.

To explicate the basic conceptual features of we-intentions and their role in some patterns of social practical reasoning we have concentrated on simple and often also somewhat idealized examples of intentional joint action. We-intentions can often have a more complex role in social practical reasoning (cf. Tuomela 1984, Ch. 7). We-intentions and the beliefs they presuppose are obviously involved in the maintenance of group cohesion as well as in group pressure against deviation from what one ought or is expected to do as a group member. They are clearly presupposed in cases of mutual support and in cases of helping behavior (when a group member seems not to be able to carry out his part of the group's task or action). Empirical evidence is required for specifying the exact role of we-intentions in this kind of various concrete cases. What we have emphasized in this paper is the intelligibility and the need of we-intentions and the beliefs they presuppose in theorizing on social action, their indispensability in social practical reasoning, and the fact that they lead to action in suitable circumstances.

Notes

- 1) On mere conceptual grounds one cannot say very much about this matter. If we, however, consider m agents A_1, \dots, A_m jointly performing a social action X , we can propose that the following principles hold true on conceptual grounds (cf. Tuomela 1984, Ch. 5):
 - (a) A_1, A_2, \dots, A_m jointly performed X if and only if $A_2, A_3, \dots, A_m, A_1$ (or any other permutation of their indices $1, \dots, m$) jointly performed X .
 - (b) If A_1, \dots, A_m jointly performed X , then $(EX_1) \dots (EX_m)$ (A_1 performed X_1 & ... & A_m performed X_m & X_1, \dots, X_m make up X).
 - (c) (i) There are actions with an upper bound (as to the number of participating agents);
 (ii) there are actions with a lower bound;
 (iii) there are, for some m , necessarily m -agent actions and thus actions with both an upper and a lower bound;
 (iv) there are actions with no lower and no upper bound.
- 2) In Tuomela (1984) a slightly different phrasing in clauses (i) and (ii) of (WI) is used. For instead of "intends to do his part of X " (clause (i)) and "do his part of X " (clause (ii)) the longer locutions "intends to do X (or his part of X)" and "do X (or his part of X)", respectively, are used. The difference, however, is verbal only. For, as will be pointed out later in the paper, in the case of some social actions such as "cleaning the house" or "picking a bucketful of berries" the

same verbal phrase in our natural language goes both for the total social action and its parts, whence the liberty of usage. The reader may, however, raise a related doubt here. For the analysandum of (WI), viz. essentially "A_i we-intends to do X", clearly involves the total social action while the analysans only speaks about its parts. Is there not missing an intentional component, so to speak? For A_i conceivably could intend to do what amounts to his part of X for some other purpose than for the participating agents' success to perform the full action X. This is, however, blocked by our understanding of the notion of a part of X. For we analyze this notion as something satisfying clause (b) of note 1) above, and thus A_i's performing his part of X can be regarded as at least conducive to X (and often also necessary for it). Accordingly, we propose that the following must hold true of the notion of a part of X: If A_i intends to do something as his part of X, then he intends that X be realized, given the truth of the analysans of (WI). Thus we claim that A_i cannot do his part of X, viz. something as his part of X, in the case when he we-intends to do X unless he does it with the purpose of the agents' jointly succeeding in performing X (with at least a nonnegligible probability).

We cannot here undertake a proper defense of what was just claimed, however. If the reader does not find the above remarks plausible he can of course reformulate the analysans of (WI) in a way which he thinks guarantees that A_i's intention, so to speak, comes to cover the whole X (even if A_i himself, of course, normally is unable to bring it about alone).

- 3) One can also have a more limited view of normal conditions, and present a more or less standard list of the *ceteris paribus* clauses (e.g., ability, exclusion of prevention and akrasia, limitations due to time etc.), leaving open the possibility that other contingent factors intervene. In this overview we don't consider the *ceteris paribus* assumptions in detail. Thus it is assumed that the participating agents have the abilities to perform their component actions, and the conditions that make the exercise of ability possible will be grouped under the notion of opportunity for action (cf. von Wright 1980, 3 ff.).
- 4) Acting in risky situations might be proposed as a counterexample to this conclusion. In these cases risk or uncertainty factors are embedded in the normal conditions, so that although the necessary X_i's are realized, the desired end is not attained. It can be argued that here we have a case where the necessary X_i's and the normal conditions are realized, but X is not realized. Thus the agents' failure cannot be due to the not-obtaining of the normal conditions, as they include the risk. An example is e.g. a team's participating in a poker game (or lottery etc.). Let us consider the case of the team's not winning, although the members did what they could, everything they considered necessary, and they had the ability to win.

One might try to answer this objection along the following lines: The members of the team we-intended to do something, but the object of their we-intentions, i.e. X, was not the joint action of winning the game. They intended to participate and wished (desired, wanted) to win. If the agents are aware of what they are going to do and of the circumstances, they cannot we-intend 'against their better knowledge'

without contradicting their we-intention. They we-intended to participate in the game as a team, and if, by chance, they win, their wish comes out true, too. If, on the other hand, they don't win, this does not contradict their we-intention to do X. One cannot separate the agents' intentions to act from what they expect will be the case (cf. Sellars 1980). Their we-intention is open, so to speak, and so is their intentional joint action, too. Their we-intention (but not their wish) to do X can be satisfied even though they don't win. What the agents we-intend to do, i.e. X, is typically to participate and to play as well as they can. If the agents are confident that they will win, if it is to be expected that they will win (they as a rule, 'normally' or always win), their we-intended action could be winning the game (and not merely participating or trying to win).

The agents, when aware of the normal conditions, do usually we-intend to try to win. This action (i.e. the object of their we-intention) is of course different from mere winning. It is accomplished when the game is over, even though they did not win. Examples where the normal conditions for this action are not satisfied would be the cases of the agents' being drugged or locked up in their rooms, forgetting about the time etc.

- 5) Sellars also emphasizes this. In his (1980) he distinguishes between state-of-affairs intentions and intentions to do. The former ("It shall be the case that ..."), ceteris paribus, imply intentions to do ("I shall do ..."). A shareable state-of-affairs intention again will, ceteris paribus, imply both I-intentions and we-intentions (I-referential and we-referential action intentions). As we-referential action intentions imply I-referential action intentions which normally imply action, in the end individual action level is reached. In general Sellars' "Shall be-intentions" are broad scenarios for the future, elaborated for choice by means of his Conjunction Introduction and So-be-it principles, they are then simplified, and finally lead to intentions to act and action (Sellars 1980, 91). Sellars also connects his analysis of we-intentions with some moral considerations. Our treatment of we-intentions is more limited and ethically neutral, but it does not contradict Sellars' analysis.

Bibliography

- Cartwright, D./A. Zander (1968), *Group Dynamics Research and Theory*, New York
- Lewis, D. (1969), *Convention: A Philosophical Study*, Cambridge, Mass.
- Sellars, W. (1980), On Reasoning about Values, in: *American Philosophical Quarterly* 17, 81-101
- Shaw, M.E. (1981), *Group Dynamics: The Psychology of Small Group Behavior*, 3rd ed., New York
- Tuomela, R. (1977), *Human Action and its Explanation: A Study on the Philosophical Foundations of Psychology*, Dordrecht
- (1984), *A Theory of Social Action*, Dordrecht

von Wright, G.H. (1980), *Freedom and Determination*, *Acta Philosophica Fennica*, 31, No. 1, Amsterdam