

Diskussion/Discussion

Marvin Belzer

Intentional Social Action and We-Intentions

Abstract: In his recent book Professor Tuomela presents a philosophical account of social action that relies upon the presuppositions of his purposive-causal theory of individual action. In particular, the concept of "we-intention" plays as central a role in the new theory as does that of intention in the earlier one. This article examines Tuomela's concept of "we-intention". Tuomela's introduction of the concept into social action theory is motivated by the assumption that theories of individual actions and social actions are analogous relative to the role of a concept of intention in those theories. This assumption is criticized; and a number of difficulties with the new concept are discussed.

Anscombe's distinction, in Intention (1957) between three main uses of the concept of intention - (i) acting intentionally, (ii) acting with an intention, and (iii) intending to act - has been important partly because it has provided a framework within which philosophers have attempted to take one of the uses as primary, and analyze the other two in terms of the primary one. For instance, in his influential early work (especially Davidson 1963), Donald Davidson took (ii), acting with an intention, as primitive - and he gave a reductive account of this, according to which a person would be said to have acted with an intention if one's action were appropriately related to one's desires and beliefs, and the intention with which one acted was to be understood as identical with the motivating desires. Nowhere in the account does Davidson posit mental states of pure intending or willing. One acts intentionally, on this account, if one acts with some intention. As for (iii), intending to act, Davidson assumed that a reductive account of this notion could be given, but he did not attempt to provide the account in his earlier writings. More recently, however, he has retreated from the reductive stance precisely because of difficulties concerning the analysis of "intending" in terms of the category "acting with an intention" (cf. Davidson 1978) and he has said recently, of his early view that "intending to act" could be understood in terms of "acting with an intention" or "acting intentionally":

"I was wrong. When I finally came to work on it, I found it the hardest of the three; contrary to my original view, it came to seem the basic notion on which the others depend ..." (Davidson 1980, XIII)

Most contemporary analytical action theorists would concur here; most theorists freely introduce unanalyzed mental episodes or attitudes of intending (one's 'having an intention') into their theories, and these mental states are regarded as unanalyzable states quite distinct from the states of desiring and believing.

Not least among such theorists is Raimo Tuomela, who in his Human Action and Its Explanation, provided an extremely detailed analysis of individual human action. One of the important contributions of the book (published in 1977) was in the way it utilized the idea that intending is a type of mental state that should be assimilated neither to states of desiring nor to states of believing. Now in his recent book, A Theory of Social Action (1984; henceforth 'ATSA'), Tuomela has attempted to extend his "purposive-causal" theory of individual human action to social actions as well.

The new book also is detailed, and it is extremely bold as well, more so than one might expect in a pioneering effort (as Tuomela's book surely is, at least for analytical philosophy of action). In the following respect it is somewhat too bold, it seems to me: Tuomela assumes from the outset that the analysis of intentional social action should take as fundamental (relative to Anscombe's three-fold distinction mentioned above) the category of "intending" ("having an intention"). The most important concept in the book is that of a "having a we-intention" which concept Tuomela explicitly defines in terms of individual intentions and beliefs. Tuomela designs the concept to play a central role in the purposive-causal theory of social action which is analogous to that played by the notion of "intending" ("having an intention") in the purposive-causal theory of individual human action.

In this discussion paper, I would like to focus on the concept of "we-intention" as Tuomela defines and uses it in his new book.

Let me say first why Tuomela's strategy strikes me as bold. It is not at all obvious, even to one who accepts the claim that an individual's 'having intentions' is necessary for the explanation of his or her intentional individual action, that one need posit special types of intentions in order to be able to explain intentional social action. Consider, for example, one type of problem for the solution of which intentions have seemed necessary at the level of individual action - that of "wayward causal chains" (cf. ATSA, 61, for a discussion of such chains) - such chains are problematic for simple 'mental cause' theories (like that found in Davidson's earlier papers) which hold generally that an action is intentional if it is caused by one's desires and beliefs. Would similar problems exist for an analysis of

intentional social action which took as fundamental, not the notion of individuals' having special types of intentions (we-intentions), but rather the notion of a group acting intentionally (or perhaps acting with an intention)? It is difficult to say, since we do not have at hand a developed theory, along those lines, relative to which we can even begin to pose the problems. At the least, however, such a theory could avail itself of the individual intentions possessed by the members of a group in order to explain what it means for a group to act intentionally (or to act with an intention).

These points are developed more fully later in this paper, for it may be that many of the problems that have been thought through in the analysis of individual's actions may have to be thought through again for social actions. Tuomela's book reflects little restraint at this point, as is particularly clear relative to the matter just discussed, but perhaps these comments do not go very far by way of criticism - for, after all, having developed a rich and fruitful theory of individual action, Tuomela scarcely can be faulted for attempting to utilize some of the fundamental assumptions of that theory in a related realm concerning which there has been very little philosophical analysis.

Tuomela borrows the term "we-intention" from Wilfrid Sellars (1963) but he gives the term a meaning of his own. In particular he differs from Sellars by claiming that the concept of "we-intention" is explicitly definable "essentially in terms of I-intentions and I-beliefs" (35), as follows:

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

(i) A_i intends to do X (or 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 X (or his part 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 X (or his part of X):

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

Before considering some difficulties concerning the 'social character' of we-intentions, there is a point concerning their logic that should be noted. The intention referred to in clause (i) of the definiens is a conditional intention. Tuomela claims that

"it can be shown to 'logically' follow from (i)-(iii) that A_i intends to do (his part of) X unconditionally." (36)

This detachment is indeed important if we are to be able to say, on the basis of (i)-(iii) that A_i intends X (or his part of X) unconditionally. However, Tuomela's claim concerning the detachment is based upon a principle like

If A_i intends that p if q and if A believes that q, then A intends that p,

concerning which many philosophers (including Tuomela himself, cf. 1977, 155, fn. 3) have remarked that it presupposes that A is rational in a strong sense. This is a shortcoming of the definition of (WI), for we can attribute we-intentions to beings only on the assumption that they are suitably rational (otherwise they may have the conditional intention of clause (i) of WI, and the belief of clause (ii), but nonetheless fail to make the detachment).²

Quite properly, Tuomela is anxious to show that even though we-intentions are defined in terms of individual intentions and beliefs, the individual intending that is involved in we-intending has a 'social character' - it has this feature, he says, because it is "conditional on the other agents' actions" (35). Certainly it is important for Tuomela's project that "we-intentions" be distinguished from ordinary "I-intentions" in some way if they are to ground a theory of intentional social action in the way that I-intentions ground his theory of intentional individual action. In fact, however, there really is nothing particularly 'social' about we-intentions by virtue of their being conditional upon the actions of others. An individual may have intentions that are conditional upon the actions of other individuals even though there unquestionably is involved neither social action nor any sort of collective intention or we-intention. Suppose Arabella intends to do X, if Barbarella does Y. Expression of Arabella's conditional intention requires reference to Barbarella, of course, but this gives Arabella's intention a 'social character' only in the superficial sense that Arabella's intending to do X, if it snows, gives her intention a 'meteorological character'. Are we-intentions supposed to have a social character only because they are conditional upon the actions of others?

Tuomela also distinguishes between intending in the "we-mode" and intending in the personal mode; for instance, on p. 121 he says "by a we-intention we mean an intention in the we-mode satisfying the definition (WI)". Perhaps this notion may be invoked to explain the social character of we-intentions. The notion of "we-mode" also is borrowed from Sellars, and while Tuomela makes only a few passing comments about the notion, he seems to accept the view that intending in the "we-mode" is to have an

intention that would be expressed in the form "We shall do X" (cf. 32). Now it is tempting to think that the distinction between modes of intending is strictly unnecessary if one simply distinguishes between two types of contents of intentions - those with first-person singular subjects and those with first-person plural subjects. Tuomela's definition (WI), however, indicates that he no longer takes the content of intentions to be a proposition (as he did in 1977, e.g. cf. 144); the content rather is "to do X" where "X" represents a subjectless action type - he apparently is assuming that the contents of intentions do not have subjects at all (and hence are not propositions or even Castañedan practitions). Unfortunately Tuomela does not address the question of content directly in the new book. Moreover, Tuomela cannot invoke the distinction between the two modes of intending in attempting to explain the social character of we-intentions; for intending in the we-mode must be reducible in some fashion to I-intending if Tuomela is to succeed (and this explicitly is his goal in formulating (WI)) in defining the concept of we-intention "essentially in terms of I-intentions and I-beliefs" (35).

The few examples employed to illustrate we-intending tend to baffle. For instance, Tuomela says for instance of the ten kilometer track race in the Moscow Olympics that "the definiens of WI may be regarded as true for every participant runner, assuming their I-intention is to run the distance successfully (or, perhaps, to win)" (37). Yet for "X" being "to run successfully the IOK race in the Moscow Olympics" it is difficult to see why Tuomela thinks the definiens of WI is true for each runner who participated. Why for instance should we suppose not only that Viren intended to run successfully, but also, in addition, that he had the complex conditional intention mentioned in clause (i) of WI? This example also underscores the importance of dealing directly with the problem of the content of intending; for if (apparently contrary to Tuomela's present view) there are subjects contained in, or associated with, the contents of intentions, then even though each runner intends to win, it is absurd to say that each of the runners intended the same thing (and hence no temptation at all to suppose that the definiens of WI would be satisfied for any of the runners). Each runner intends to do the same thing (i.e., to win) but the content of the intending is different in each case.

There are a number of further difficulties concerning the social character of we-intentions. Clause (iii) of WI specifies that a necessary condition for A_i 's we-intending to do X is that each member of G believes (i) and (ii) of A_i , and believes that each member believes (i) and (ii) of A_i etc. (cf. the definition of "it is mutually believed in G that p", 206). This is the meaning of clause (iii) as stated, but Tuomela seems to give the clause a much different interpretation, according to which each member A_j of G must believe of every member A_k of G (and not merely of A_i) that conditions (i) and (ii) are met, and also believe that each member believes

(i) and (ii) of A_k , etc. That Tuomela means something like this is indicated by his comment that the existence of we-intentions to do X "entails mutual belief in everyone's intending to do (his part of) X" (125).³

A puzzling feature of WI, on either reading of clause (iii), is that it is not difficult to imagine situations in which conditions (i)-(iii) are met, and yet A_i is the only member of G for which conditions (i) and (ii) hold. This means that one can we-intend to do his part of X, relative to G, even though one is the only member of G who intends to do his part of X. Suppose that at a certain time Arabella intends to do her part of moving the piano, given that a sufficient number of others do their parts, and that she also believes that a sufficient number of others will do their parts; but there does not exist 'mutual belief' to that effect (on either reading of clause (iii)). If not, Arabella does not have a we-intention. If the mutual belief later comes into existence, even if nobody else intends to do her part of moving the piano, then Arabella's intention becomes a we-intention. One wonders, however, just why the beliefs of the other members of G are either necessary or, in this example, sufficient, for Arabella's we-intending.

Another difficulty with WI is that clause (ii) makes it appear that one has we-intentions in many action situations only if one's acting according to the intention is positively irrational. Clause (ii) stipulates that one believes of one's group that a sufficient number of its members, as required for the performance of X, will do their parts of X. This means of course that one believes that the group's doing X will occur. Arabella can we-intend to do her part of moving the piano only if she believes that sufficiently many members of her group will participate, so as actually to move the piano. If she believes this, why should she intend to participate at all? The question is important, even supposing that Arabella wants very badly for the piano to be moved. For, if (as is often the case in social actions), Arabella's helping out requires some sacrifice of her, then the belief required by condition (ii) surely would be a reason for not intending to participate at all.⁴ On the other hand, why cannot one intend to do his part of a social action, even though he is pessimistic about others doing their parts? Why would it not be sufficient, for instance, that one desired or intended that others do their parts? It may be replied that in such cases there is mere I-intending, not we-intending; yet it is not clear why the notion should be so restricted. Even if the reply were accepted, at most condition (ii) should require merely the belief that sufficiently many members, for the performance of X, intend to do their parts, rather than the belief that they will do their parts.

An analysis of intentional social action must give some account of a group's having a common goal or intention. It is important for us to consider

whether Tuomela's notion of we-intention can be used to characterize the notion of shared goals within a group. Tuomela in fact does assimilate the notions of we-intention and group-intention. The discussion of the definition WI on p. 38, for instance, moves almost indiscernibly from speaking of an individual's we-intending to do X to talk about "a group's (viz. our) we-intending to do X", and it is held that in intentional joint action the relevant group intention expresses the agents' shared common goal (118-9). We-intentions are expressed by statements of the form "We shall do X". The notion of we-intention would be of great value if it could be used to elucidate the notion of shared goals. There are however additional problems concerning the social character of we-intentions which make it doubtful that the concept of we-intention can be used to shed light here.

Notice first that even if each of the members of a group we-intends to do X, this does not mean that there is anything at all that is 'shared' by the members of the group. The reason for this, oddly enough, is that according to WI, A_i may have a we-intention to do X even though she has no actual mental state of any type that would be expressed by "We shall do X". To see this, suppose that condition (i) of WI is satisfied because A_i intends conditionally to do her part of X, given that sufficiently many others do their parts. Because each member may be in a similar position, it follows that each member may have a we-intention to do X even though nobody at all has an intention that would be expressed by "We shall do X". If Arabella conditionally intends merely to do her part of moving the piano (if for instance she intends to carry the piano bench), while Barbarella conditionally intends merely to do her part (holding the door open) then the conditions (i)-(iii) of WI may be satisfied so that each may be said to have the we-intention to move the piano - even though neither of the individual intentions involves any reference at all to this action type (indeed, even if neither has any thought at all about our moving the piano). Such cases may not be 'fully intentional' social actions, but that does not affect the main point here - that, as we-intentions are defined in WI, members may have the we-intention to do X even though nothing at all like a 'shared goal' exists.⁵

The notion of we-intention is a theoretical construct that is only very loosely tied to psychological and social reality. As noted above, Tuomela grounds his theory of intentional social action on the primacy of "intending to act" - hence the importance of the concept of we-intention. The unsuccessful attempt to assimilate shared common goals to the notion of we-intention indicates that some alternative starting points should be considered. The notion of we-intending itself could be defined in terms of shared common goals - that is, the notion of a group's having a shared common goal, or its having an intention, might be taken as primary. Here is a start: a group has the intention to do X if sufficiently many of its

members for the performance of X have individual intentions expressed in the form "We shall do X"; and an individual member of a group has a we-intention to do X if the group has the intention to do X. We-intentions are defined in terms of individual intentions, as in Tuomela's theory; and even though according to the alternative one can 'inherit' we-intentions from one's group (a result that seems appropriate), the alternative is tied much more closely to social reality and (because WI permits the we-intention to do X even if one has never had an explicit thought about Xing) it is tied no less closely to the individual's own mental states than is WI.

Before developing alternative conceptions of we-intending, however, I believe it is important to consider the question whether or not it is realistic to hope that an analysis of intentional social action can be grounded successfully on the notion of a group's intending to act, especially when as for Tuomela holistic group entities are abjured. Intending, or having an intention, is a state of mind. There is widespread consensus that such states are fundamental for the analysis of individual human action. As was pointed out at the beginning of this paper, however, it may be that such states are fundamental for the analysis of social action only because the concept of individual intending may play a role in formulating the fundamental concepts in a theory of social action. From the fact that individual intending is conceptually fundamental for the theory of individual action it simply does not follow that group-intending is to be basic for a theory of social action, much less that an individual's 'we-intending' is fundamental.

Tuomela defines a collective's acting intentionally in terms of a group's acting with appropriate we-intentions (145-6). Instead of postulating we-intentions to account for intentional social action, however, it seems to me that (relative to Anscombe's three-fold distinction) it is quite likely that none of the three categories should be regarded as 'more basic' than the others, but rather that each of the notions of (i) a group's acting intentionally, (ii) a group's acting with an intention, and (iii) a group's intending to act should be analyzed directly in terms of the intentions and actions of individuals in the context of a set of plans, policies, and primary and secondary rules, accepted by the group, for assigning roles and responsibility. The 'multi-agent' actions of groups that are not explicitly organized, as well as the actions of organized collectives and institutions, may be understood along these lines by appeal to social conventions some of which, like secondary rules of law, permit individuals to formulate and accept relevant rules by their intentional action in a social situation.

Notes

- 1) The particular detachment principle presupposed by Tuomela in ATSA would be like this one except "that p" would be replaced by "to do X".
- 2) The assumption concerning detachment is awkward for the additional reason that formal representation of expressions of conditional intentions may require the use of a subjunctive conditional for which ordinary modus ponens fails - so that even a 'strongly rational' being of whom clauses (i) and (ii) were true would not necessarily make the detachment presupposed in WI (as might be the case were there some Q such that she also believes that Q and intends not to do her part of X given that she believes both that a sufficient number of members of G will do their parts of X and Q). I cannot address these complexities here and wish only to note that the concept of we-intention as defined depends heavily upon a controversial view concerning the logic of expressions of conditional intention.
- 3) That this interpretation of clause (iii) is intended also is evidenced by Tuomela's comments concerning the question whether a person's we-intending to do his part of X entails each person's we-intending to do his part of X. On p. 36, Tuomela says

"The definiens of (WI) entails that each member of G intends to do either X or his part of X."

and this makes sense only on the second interpretation of (iii), as does the contradictory claim made on p. 38 that, for person M in group G,

"M's we-intending does not entail that any other member of G I-intends, for the mutual belief in our definiens need not be true."

(On either reading of clause (iii) it is clear that Tuomela misspoke himself in the earlier passage on page 36.)
- 4) Those situations involving the 'free rider' problem are such situations. If Arabella believes that sufficiently many other members of her group will work to move the piano, then obviously her contributing is not necessary for the good to occur.
- 5) Condition (i) of WI may be read either de dicto or de re. The discussion here presupposes a de re reading, so that Arabella may intend to do her part of moving the piano by virtue of intending conditionally to carry the piano bench in circumstances in which her carrying the piano is in fact 'her part' of moving the piano. Tuomela might wish to strengthen (i) so that a de dicto reading is required: to we-intend to move the piano, Arabella would have to intend not merely to carry the piano bench in the circumstances described but she also would have to have an intention explicitly to do her part. The new condition is extremely strong. Even the strengthened condition, however, does not guarantee that, when the members have identical we-intentions to move the piano, there also is a shared common goal expressed e.g. by "We shall move the piano". For even the strengthened definiens of WI could be satisfied even though there exists nothing more than similar individual intentions each to do his or her part of the

action. These points also are important for the "purposive-causal" theory of social action. The patterns of practical inference which play an essential role in that theory contain key premises of the form "We shall do X", cf. (S), p. 121, and (W2), p. 34. Effective intendings are the individual intendings that result in action, and where X_i is my part of the social action X, my effective I-intending is expressed by "I will do X_i ". The important point is this: effective I-intendings are to be "purposively obtained" from effective we-intendings, expressed by "We shall do X", as e.g. in pattern (S), cf. p. 122. Yet this "purposive-causal" role of we-intendings is undermined by the fact that one can have we-intendings by virtue of having the effective I-intending, and without accepting anything at all like the claim "We shall do X".

Bibliography

Anscombe, G.E.M. (1957), *Intention*, Oxford

Davidson, D. (1963), *Actions, Reasons, and Causes*, in: *Journal of Philosophy* 60, 685-700

- (1978), *Intending*, in: Y. Yovel (ed.), *Philosophy of History and Action*, Dordrecht

- (1980), *Essays on Actions and Events*, Oxford

Sellars, W. (1963), *Science, Perception and Reality*, London

Tuomela, R. (1977), *Human Action and its Explanation: A Study on the Philosophical Foundations of Psychology*, Dordrecht

- (1984), *A Theory of Social Action*, Dordrecht-Boston-Lancaster