

Raimo Tuomela · Maj Tuomela

Cooperation and trust in group context

Received: 9 June 2004 / Accepted: 21 October 2004
© Fondazione Rosselli 2005

Abstract This paper is mainly about cooperation as a collective action in a group context (acting in a position or participating in the performance of a group task, etc.), although the assumption of the presence of a group context is not made in all parts of the paper. The paper clarifies what acting as a group member involves, and it analytically characterizes the “we-mode” (thinking and acting as a group member) and the “I-mode” (thinking and acting as a private person).

The main task of the paper is to clarify cooperation, and this is done by dividing cooperation into I-mode and we-mode cooperation. I-mode cooperation relies on the participants’ relevantly adjusting their goals toward others’ goals and actions to the benefit of all participants. We-mode cooperation basically amounts to joint action in the we-mode. These kinds of cooperation are also connected to the respective kinds of acting as a group member.

The paper also investigates the relationship between cooperation and trust and, basically, argues that a two-way connection is involved.

Keywords Cooperation · I-mode · We-mode · Group membership · Trust · Respect

1 Introduction

It is a platitude that people can act privately or as members of groups. All of us belong to many groups and we spend much time acting as group mem-

bers. In our daily life we act as members of some social group – a factory, a business firm, the government, the school, the family, or some other of the myriad of groups. In this paper our focus will be on groups (or, what below amounts to the same, collectives) which are externally and internally autonomous, viz. groups that can determine their own goals, standards, beliefs, etc. free from group-external coercion (external autonomy) and in which the members can in principle act freely and e.g. collectively decide about group matters (internal autonomy). Our central aim is to investigate how acting as a group member relates to cooperation and how acting as a group member and cooperation relate to people trusting each other when so acting. More specifically, our basic questions in this paper are:

- (1) What kinds of cooperative collective action are there?
- (2) What is cooperation in a group context?
- (3) Does cooperative collective action require trust, and does cooperation enhance trust?

Cooperation is a complex notion and we cannot here much discuss the notion itself. Basically, cooperation is collective acting towards a shared goal, but this can be understood in several different ways. We will not discuss such problematic questions how cooperation in collective action dilemmas can be achieved or under what conditions it is rational or plausible to cooperate with some other persons. The work of R. Tuomela has previously been applied to give plausible answers to them (Tuomela 2000). Here we will be mainly investigating contexts where people act as group members and are required to cooperate by the group, although our basic classification of cooperative collective action also covers other, “non-group” cases. In the present paper we will connect cooperation and acting as a group member and show how trust and cooperation are related.

The plan of the paper is as follows. Section 2 gives a survey of the various ways in which one can act as a group member. It gives a summary of the topic discussed in detail in Tuomela and Tuomela (2003). Section 2 also summarizes the we-mode/I-mode distinction, drawing on R. Tuomela (2002a,b), and Tuomela and Tuomela (2003). Section 3 presents our lengthy classification of cooperative collective action that covers both intentional and non-intentional cooperation and cooperation as a group member and as a private person. Section 4 applies the conceptual machinery of the paper to the question of whether cooperation needs trust and whether cooperation promotes trust (see M. Tuomela 2002, 2003a). There is also a concluding summary (Sect. 5).

2 Acting as a group member and the I-mode/We-mode distinction

2.1 Acting as a group member

Human beings are social beings, and it has been argued that acting as a group member is an innate property of humans, as they have evolved in a group context: they are genetically disposed to act as group members and to

cooperate with other group members (cf., Richerson and Boyd 2001, 2004, for arguments). Accordingly, it is reasonable to suggest that also the dispositions to think and act in the we-mode (in a rudimentary sense, at least) and hence to be in collectively intentional states are genetically based, although the contents of we-mode thoughts and actions are of course culturally and socially determined. There are of course also various other factual reasons, e.g. there are central tasks factually impossible to perform without cooperative activities and, on the other hand, people often cooperate on purely social grounds, because of their inherent sociality (see Tuomela 2000, Chapt. 11, for a long list of reasons for rational cooperation).

In the present section we will concentrate on the question of what acting as a group member is and give precise summary accounts of some of our notions of acting as a group member and also the I-mode and we-mode concepts that we need in this paper. (The reader is referred especially to R. Tuomela 2002a,b, and Tuomela and Tuomela 2003 for a comprehensive treatment and extensive discussion.)

Our account relies on a kind of public-private distinction in that it distinguishes between acting in a group context from acting in a private context. When people act as group members this can only take place by their acting in the socially and normatively right circumstances, viz. in the right group context. For instance, acting in a group position (task) and thus performing one's work duties requires that such circumstances, partially explicating the group context, be present (cf., Tuomela 1995, pp. 234–241, for discussion). In any group it is possible to perform freely chosen actions qua a group member provided that these actions – or, more broadly, activities, including mental ones – are within the “realm of concern” of the group, viz. provided that they belong to topics which are of concern or are of significance for the group and that they are performed in the context of the group. Such actions are to be (rationally) socially accepted by the group (either through normative, group-binding group acceptance or through the we-acceptances by the group members or their majority, cf., Tuomela 1995, Chapt. 7). These actions and attitudes belong to the topics that a group, *g*, is concerned with in group contexts. The set of all such topics is the realm of concern of the group.

The subset of the topics (expanded to include also e.g. attitude contents over and above action types) that the group has accepted as to express its *constitutive* goals, values, standards, beliefs, norms, etc. is called the “*ethos*” of the group. If a group has no constitutive goals, etc. it must at least have some central goals, beliefs, etc. that in such a case will constitute its ethos. The ethos serves to constitute the group by giving the purported content of its ontological identity. The notion of ethos also characterizes part of the notion of a group context (as contrasted with a private context) and serves to specify the actions that qualify as actions as a group member. As a consequence it becomes true that an agent performs an action as a group member in the fullest sense if and only if it is collectively accepted by (or is collectively acceptable to) the group members as an action that promotes, or is at least weakly conducive to, the ethos of the group.

We will assume as a simplification here that the realm of concern of the group is given and it conceptually precedes the ethos of the group. This

assumption makes it easier to account for the group's changing its ethos and identity and for comparing different (and perhaps competing) groups concerned with the same kinds of topics. Furthermore, the notion of acting in the realm of concern is required for elucidating acting as a group member below. Our ultimate conceptual explicans here is (rational) collective acceptance (typically in the we-mode sense).¹ The realm of concern, the ethos, and acting as a group member all ultimately depend on collective acceptance.

The general case to be considered is that of a normatively structured, externally and internally autonomous social group with positions (such as professor, secretary, assistant, characterized in terms of a task-right system in the sense of Tuomela 1995, Chapt. 1), while the unstructured case can be regarded as its special case with no specific positions over and above mere group membership. We will first classify the types of actions within the action realm of a (structured) group's concern. These actions will all be collectively acceptable and mutually knowable as being at least weakly ethos-promoting. Here are the four action classes (cf., R. Tuomela 2002a, b, and Tuomela and Tuomela 2003, especially pp. 15–16):

- [1] Positional actions (related to a norm-based group position or role based on normative group expectations), which include (i) actions (tasks) that the position holder in question *ought to* perform, perhaps in a special way, in certain circumstances and (ii) actions that he *may* (is permitted to) perform in some circumstances;
- [2] actions which constitutive and other general group norms and group standards (e.g., norms and standards which are not position-specific) require or allow;
- [3] actions and joint actions that are based on situational intention formation (e.g., agreement making) that has not been codified in the task-right system of *g* or the group norms of *g* but which still are consistent with actions in [1] and [2] and are ethos-promoting;
- [4] freely chosen actions or activities (and possibly joint actions), which are not incompatible with actions in classes [1]–[3]; these freely chosen actions belong to the realm of concern of *g* and are rationally or reasonably collectively accepted by, or acceptable to, the members of *g* as ethos-promoting actions.

Acting as a group member – in the *core* sense – is to act intentionally within the group's realm of concern, promoting or at least respecting the ethos, viz. the central, constitutive goals, values, standards, beliefs, norms, elements of the history of the group. The ethos is supposed to give the

¹ Collective acceptance in the meant fullest sense can be argued to amount to coming to hold and holding a relevant shared we-attitude in the we-mode, where the we-attitude is either a we-intention or a we-belief or, slightly more broadly, a we-attitude with either a world-to-mind or mind-to-world direction of fit of satisfaction, as the case may be. Thus, when we say that e.g. group values are collectively accepted this means that the members are collectively committed to furthering, with the world-to-mind direction of fit of satisfaction, certain values (that will amount to group values in the account) for the benefit of the group. (See R. Tuomela 2002a, Chapt. 5, for discussion.)

members their main reason for action when acting as group members. The core sense can in a general way be divided into “we-mode” acting and “I-mode” acting as a group member. (However, strictly, speaking, the respective precise definitions (*WM*) and (*IMI*) of Subsect. 2.2. need not yet be satisfied; analogously, groups can analogously be divided into we-mode and I-mode groups on the basis of their members’ typical kind of action.) The former kind will officially be called the *standard* and the latter the *weak* sense of acting as a group member. Both are needed for accounting for group life. The basic situation is this: if there is collective acceptance of the ethos, it will be upheld and maintained by the participants by means of actions falling within [1]–[4] (the ethos has the world-to-mind direction of fit of semantic satisfaction). In this sense the ethos is taken by the members to be the group’s ethos and to be constitutive of the group. If they *jointly* intend to maintain and satisfy the ethos, there is collective commitment to it, where collective commitment means that they have collectively, as a group, bound themselves to the content of the joint intention and are socially committed to each other to be so bound.² If they only separately intend to maintain and satisfy it, we get separate intentions and accordingly private commitment (viz. commitment as a private person, not as a member of the group although a member of the group) to the ethos.

In classes [2], [3], and [4] the actions may also be based on the group’s *non-constitutive* goals (etc.), thus goals that are not of the kind central to the group’s identity. Still the actions will at least in an indirect sense be taken to be ethos-promoting. Actions in class [1] are typical positional actions, and subclass (ii) of [1] consists of actions that the holder of a position may choose from. (The task-right system specifying (i) and (ii) may contain rule norms and/or proper social norms in the sense of Tuomela and Bonnevier-Tuomela 1992, and in improved form in Tuomela 1995, Chapt. 1.) However, actions in classes [2]–[4] can occur also in the positional case. Note that in the case of unstructured groups, class [1] can be empty, when a strict notion of normative position is meant.

Action as a group member can be either successful or unsuccessful. There may thus be failures due to false beliefs about the group’s norms and standards or what obeying the ethos in concrete situations amounts to, as well as failures due to lack of skill, or due to environmental obstacles. What is required is that the group member will intentionally attempt to act in a way that he takes to be within the group’s realm of concern without violating the group’s ethos. The agents share the ethos and are committed to it and therefore ought to promote it – or it should be at least a “presupposition-stating” reason for the member’s action. Thus, it need not be a salient motivational reason, as especially classes [3] and [4] indicate. The notion of ethos can be understood in a wide sense in which every group can be taken to

² In a simple two-person case, you and I are collectively committed to an item *p* if and only if you and I are both committed to *p* (e.g. intend or believe *p*, as the case may be) and we mutually know this and our shared commitment satisfies the “Collectivity Principle” (see R. Tuomela 2002a, Chapt. 2). This principle here amounts to saying the following: necessarily due to or collective acceptance to intend or believe (etc.) *p* in our dyad, your commitment is satisfied if and only if my commitment is satisfied

have an ethos, which may consist just of some basic shared ends or beliefs or something closely analogous.

Acting as a group member in the *core* sense is equivalent to acting intentionally in an “ethos-promoting” sense in terms of performing actions falling into one of the classes [1]–[4] – or at least attempting so to act. (One cannot non-intentionally act as a group member, although one can non-intentionally conform to the ethos.) The actions in classes [1]–[4] need not (especially in case [4]) be strictly contained in, or specified by, the ethos; it suffices for our present purposes that they at least weakly promote the maintenance and satisfaction of the ethos and are consistent with it. That they promote the ethos is at least a partial motivational reason for these intentionally performed actions.

Here are more exact definitions of our two basic notions related to a group, *g*:

- (a) An action *X* performed by some member or members of *g* is an *action as a group member in a weak sense* if and only if *X* is performed in the group context in the core sense (*X* belonging to one of the action classes [1]–[4]) in part for the reason that the members are privately committed to the shared ethos, *E*, of *g*, which they hence ought to promote. Here “private” is used to denote thoughts and activities that a person only personally is engaged in rather than as a full-blown member of a group.
- (b) An action *X* performed by some member or members (or authorized representatives) of group *g* is an *action as a member of g in the standard sense* if and only if (i) either the members of *g* rationally (reasonably) collectively accept for the group with collective commitment that *X* is an action in one of the senses [1]–[3] (ii) or the action *X* falls in class [4] (and is thus reasonably collectively acceptable as an action within the realm of concern of *g* and indeed as an action as a group member), *X* here being assumed to be performed in part for the reason that the members are collectively committed to the ethos, *E*, of *g*, which they hence ought to promote.

In general, the group members are obligated to perform actions within [1]–[4] and actions that accordingly are required for their performance or are otherwise ethos-promoting (these latter kinds of actions can be regarded as falling within class [4]). Actions in a group context which are not even weakly promotive of the ethos do not count as actions as a group member in the standard or weak sense³.

Notice finally that having an attitude as a group member can in the obvious way be analyzed in terms of four classes of attitudes analogous to the four action classes discussed above.

³ Because of the group commitment, the group is at least to some extent responsible for group members’ violations of the ethos. Such actions will count as actions as group member in a certain sense that we have called the “dissidence” sense. For relevant discussion of group responsibility, see Mäkelä and Tuomela (2002), M. Tuomela (2003a), and Tuomela and Tuomela (2003)

2.2 The I-mode and the we-mode

We can say that the we-mode is the mode of the *group perspective*. The notion of group perspective gets a central part of its content from the notion of *forgroupness* (or “*for-us-ness*”), which concerns the promotion of the ethos (cf., “We intend to uphold E for our group by means of actions in classes [1]–[4]”, “We believe that p for the use of our group”). The other central element is *collective commitment*, which, especially the *social commitment* involved, dynamically “glues” the members to each other and to the ethos of the group.

In our analyses in the present section the notion of group can be taken to be a weak one that does *not* entail that the group can act (either in an objective or in an inter-subjective sense of “can”), contrary to the central cases in the other sections of this paper. However, the members of the group must mutually believe (either in a *de dicto* sense, viz. under a certain description, or in a *de re* sense, viz. in the sense of direct “acquaintance” with, if not all members, at least some typical ones) that they are its members, and they are assumed to have at least approximate mutual knowledge of its ethos (its constitutive features). The group is “we” for them, but this need not be a “conative” we, one leading to collective activities, but may be only a we relating to jointly experiencing things of common concern (cf., Schutz 1967).

The group, *g*, in question can be either a normatively structured or an unstructured one. In the analyses below we are dealing with a group member who acts as a group member in a relevant sense and we next present some central modes in which he could have an attitude (ATT):

(WM) Agent A, a member of group *g*, has a certain attitude ATT with content *p* in the *we-mode* relative to group *g* in a certain situation *C* if and only if A has ATT with content *p* in part because (i) he is functioning (viz. experiencing, thinking, and/or acting) as a member of *g* in the standard sense and because (ii) ATT(*p*) (the attitude with its content) has been collectively accepted under conditions of mutual belief in *g* as *g*’s ethos-promoting attitude in *C*, or A qua functioning as a group member in standard sense has accepted ATT(*p*) as his ethos-promoting attitude in *C*, ATT(*p*) being at least collectively acceptable as ethos-promoting in *g*; and hence (iii) because of (ii), A is collectively committed in the ATT-way to content *p* at least in part for *g* (viz. for the benefit and use of *g*) in *C*.

(IM1) Agent A, a member of group *g*, has a certain attitude ATT with content *p* in the *plain I-mode* relative to *g* in a certain situation *C* if and only if A has ATT with content *p* in part because he is functioning as a group member of *g* in sense (a) (the weakest sense of acting as a group member) and is hence privately committed in the ATT-way to content *p* at least in part for himself in *C*.

(IM2) Agent A, a member of group *g*, has a certain attitude ATT with content *p* in the *pro-group I-mode* relative to *g* in a certain situation *C* if and only if A has ATT with content *p* in part because he is functioning as a member of *g* in sense (a), and is privately committed in the ATT-way to content *p* at least in part for *g* in *C*.

While obviously we-mode acting in the sense of (*WM*) goes together with acting as a group member in the standard sense (and, as will be seen, we-mode cooperation), (*IM2*) goes together with acting as a group member in the weak sense (and with I-mode cooperation).⁴ Let us note that if in (*IMI*) no relativity to a group is present (and *g* is deleted or is the “empty” group) we arrive at the pure private mode: A has a certain attitude *ATT* with content *p* in the *pure private mode* in a certain situation *C* if and only if A has *ATT* and is privately *ATT*-committed to content *p* only for himself in *C*. The pure private mode can be treated as a degenerate special case of the I-mode.

Action modes can now be accounted for by means of attitude modes and the “because of” relation. The term “because” in general expresses both reason and cause in the present kind of context and is thus both a causative and a teleological notion: an action is performed in a certain kind of mode if and only if it is performed because of (or on the basis of) an attitude had in that same mode. Furthermore, we can take the psychologically effective reasons for action to be contents of attitudes or, in some special cases, the attitudes themselves. Then the present account of an action performed because of an attitude in a certain mode amounts to saying that this action is performed for the reason expressed by the content of that attitude or, in some cases, that attitude with its content such that the reason is in the mode that the attitude is in (cf., Tuomela 1984,1995) for “because” as expressing “purposive causation”). In this sense we can speak of e.g., we-mode and I-mode reasons for action. As we will see, the simpler kinds of cooperation not based on full-blown joint action will involve I-mode reasons, while proper joint action, which indeed constitutively is cooperative, is action for a we-mode reason.

Due to ethos-related forgroupness (viz. the aspect that the content of a we-attitude is meant for the use and benefit of the group) the reason for a person’s performing an action *X* (of the right kind) as a group member in the we-mode cases is “*X* is what the group’s ethos requires or what serves it”. When the members collectively accept such contents for the group they can use them in their inferences and act upon them. In the I-mode cases forgroupness reason is also possible (and even required in case (*IM2*)) and amounts to advancing the group’s goals and interests.

Above in (*WM*) we assume that agent *A* is we-committed in the we-mode to content *p* (a content within the group’s realm of concern) in the *ATT* way just in case he participates in the collective commitment in the way that is special to the attitude *ATT* in question, the special way meaning the right kind of direction of fit of semantic satisfaction. Thus in the case of an intention, for instance, he is supposed to respect the direction of fit of the satisfaction condition of intention, viz. world-to-mind direction of fit, while in the case of belief he takes into account the fact that a belief typically has the mind-to-world direction of fit.

⁴ One can distinguish between the I-mode and the private mode as in R. Tuomela (2002b) by saying that on an occasion one functions in the private mode if one does not only not function as a member of group *g* but in addition not as the member any group. In this paper, however, we for simplicity’s sake speak of the private mode and the I-mode interchangeably

It is not possible here to enter a more detailed discussion of the I-mode and the we-mode, for that see R. Tuomela (2002b) and Tuomela and Tuomela (2003).

Here are some obvious further points related to our analyses that we state without further comment (we draw on R. Tuomela 2002b, where a more detailed discussion is available):

- (i) The I-mode (in either of the senses (*IM1*) or (*IM2*)) does not entail the we-mode.
- (ii) The we-mode does not entail the I-mode (in sense (*IM1*) or (*IM2*)).
- (iii) Social activities such as group actions and social practices may involve mixtures of we-mode and I-mode actions.

3 Types of cooperative collective action

3.1 I-mode collective cooperative action

As seen, I-mode thinking and acting is just purely personal (or “private”) thinking and acting, without the use of a full-blown group perspective. In contrast, we-mode thinking and acting means thinking and acting in congruence with a group’s ethos, viz. its constitutive goals, values, standards, norms, beliefs, etc., with the aim to satisfy and uphold them.

To begin our discussion of the nature of cooperation, we will here only be concerned with intentional cooperation below (one can even claim that all proper cooperation is necessarily intentional). Accordingly, it is assumed that the participants must *intend to cooperate*, respectively, in the I-mode or in the we-mode sense – possibly under their own conceptualizations of these kinds of cooperation (cf., below).

Cooperation has two central dimensions:

- [1] Its teleological content (viz. its end or *telos*, what the participants intentionally strive to achieve),

and

- [2] its social content, viz. interaction with some other agents relative to the content in question.

[1] gives *teleological intentionality* (“directedness” intentionality). This dimension must always exist, even in cases in which the agents act with no further goal of theirs except cooperation itself (e.g., I perform an action with no further goal but still by means of this action purport to adjust my action to your action and thus to cooperate with you). In feature [2] we are dealing with *intentionality in the social connectedness* sense. This case deals with social “aboutness” intentionality and contrasts with the “directedness” intentionality of [1]), and it consists in the agent’s epistemic and normative relationships to the participants relevant to cooperation. It need not involve more than treating the other agents as “intelligent parts” of one’s environ-

ment but it will involve the agent's intention to cooperate (at least in the reciprocity or "adjustment of action" sense – see below).

To discuss I-mode action, in such collective cooperative action the actors cooperate individually intentionally but not collectively (or, equivalently, jointly) intentionally. In this case a participant must believe that he is cooperating with somebody or some others and that they are cooperating with him, although he may not know or even have specific beliefs about who these other participants are. Furthermore, his beliefs concerning the others' identities and participation may be false. There can be full or less full knowledge of the social connectedness relating the participants, but the agent must minimally have a belief to the effect that there are other cooperating participants.

We shall now give a stylized account of I-mode cooperation in a simple two-person case where each participant performs one or more actions and has one central goal in the situation in question. The participants' intentions to cooperate entail that their means actions (viz. their actions that they use as the means to achieve their goals) are either naturally or institutionally interdependent in the situation in question, or that these actions simply are voluntarily made interdependent by the participants due to their forming the intention to cooperate. We argue that the participants must, conversely, also regard their goals (viz. the satisfaction of their goals) as interdependent due to the dependence of their means actions when cooperating. The goals may, in addition, be *antecedently* dependent and may even coincide. If rational, the participants must mutually believe that they are dependent. In all, in I-mode cooperation both the means actions and the goals of the participants are interdependent, under conditions of mutual belief.

I-mode cooperation essentially is adjusting one's means actions and goal (assuming only one is at stake here) to the other participant's actions and goal so as to further both the other's goal (viz. goal satisfaction) and one's own goal, the latter possibly only through the other's furthering it by his actions. As indicated, the means actions may be naturally or institutionally dependent prior to action, and in some cases the agents' connected goals will make them dependent even when there is no prior action dependence. Thus, for instance, we may cooperate in the I-mode so that my goal to go to Paris to take care of my business is connected to your goal to go to Rome to take care of your business by my adding my promotion of your business while in Paris to my antecedent goal and by your adding your promotion of my business while in Rome to your antecedent goal. Even if there is no behavioral dependence between the means actions, the latter still have become dependent – your means action will further the satisfaction of my goal, and vice versa. (If the agents here formed a joint intention to perform the cooperative joint action to satisfy the joint goal made up of each agent's goal, the same example works also in the we-mode case – to be discussed in next Subsect. (3.2). Then each agent performs his part of the joint action as his part.)

Our account is modeled but improves on the account given in Tuomela (2000), Chapt. 1, where also the expression "reciprocity cooperation" is used for I-mode cooperation (a cooperative kind of "I-mode joint action") and where the adjustment of actions is justified in terms of shifting one's preferences towards the other's preferences. The participants need not be assumed

to cooperate in a group context, thus they may just cooperate in the degenerate sense of the I-mode that was called the purely private I-mode in the previous section. Here is our proposal, the analysans of which can be seen to entail other-regarding (and in group context, pro-group) I-mode action:

(CIM) A_1 and A_2 (*intentionally*) cooperate in the I-mode with respect to their (proximate or further) goals G_1 and G_2
if and only if

- [1] A_1 and A_2 have the goals G_1 and G_2 , respectively, as their intended private goals (viz. goals as private persons in contrast to goals as full-blown members of a social group);
- [2] they intentionally perform respective dependent means actions x_1 and x_2 believed by them to be (at least indirectly, via the other's action) conducive to their respective goals so that each takes himself thereby to have adjusted his acting and goal to the other's action and goal with the purported result that the other's achievement of his (possibly adjusted) goal is furthered and that, by the other's analogous acting, also the achievement of one's own (possibly adjusted) goal is furthered.

Note that the participants need not know what the other's goal is, although he must believe that the latter has one in the present context. To get a notion of (functionally) *rational* I-mode cooperation from the above conceptually minimal notion, we require true beliefs in [2] and add two clauses saying that the information in clauses [1] and [2] must be mutually believed (clause [3]) such that the mutual belief is not idle (clause [4]):

- [3] A_1 and A_2 rationally mutually believe that [1], [2].
- [4] [2] in part because of [3].

We will comment on [3] and [4] later and here start by discussing clauses [1] and [2]. Our account assumes that both the means actions and the goal of the participants have, in general, been adjusted to take into account the other's respective means actions and goal.

Some comments on the variety of dependencies that might be involved in [2] are now due. First we note that G_1 and G_2 may – e.g. in cases of spontaneous cooperation with no further individual goals – be just the proximate goals conceptually inherent in intentional actions, viz. the so-called result-events of x_1 and x_2 , respectively. We already noted that action dependence might be due to antecedent goal dependence (although, conversely, some amount of indirect goal dependence anyhow comes about due to means action dependence). For example, we might have $G_1 = G_2$ such that these goals can only be satisfied by the same state token. Then the means actions cannot rationally conflict, and A_1 's action must contribute to A_2 's goal, and vice versa. A case in point is where two drivers coming from opposite directions get out of their cars to remove a fallen tree from the road; their shared goal is $G_1 = G_2 =$ the log is removed from the road. The removal of the log might take place in terms of I-mode cooperation in the sense of (CIM).

However, for some situations sharper and stronger formulations of action dependence may be used. Thus, thinking of a game-theoretic interaction case,

one may require optimal satisfaction of a participant's goal, entailing that he has to choose his action x_1 so that it is optimal for his goal in the sense of a suitable criterion (be it maximization of expected utility or something weaker) while at the same time facilitating the other's reaching his goal. A still stronger dependence requirement is that, relative to goal satisfaction, x_1 be (optimally) performed if and only if x_2 is performed, and conversely. Here the goals may be only parallel (be they the same or different) and satisfied by different event (or state) tokens or they may be the same in the stronger sense of being satisfiable only by the same token, in a sense divided by the participants through their actions. An example of the parallel (but dependent) goal case is offered by a block world case where A_1 is not able to move one of his blocks to his goal location unless A_2 first suitably moves one of his blocks, and vice versa (cf., Rosenschein and Zlotkin 1994, for this kind of cases).

Our account can make room for completely altruistic and completely selfish cooperation. In the case of pure altruism, a participant, by taking "costs" exceeding his gain from his achieving his antecedent goal, acts in order to further the other's goal and, indeed, in effect makes it his goal to satisfy the latter's goal and takes the existence of his "new" goal to depend on the existence of the latter's goal. In the case of purely selfish cooperation the participant performs his action by which he furthers the other's goal purely as a means for furthering his own goal. Depending on the goal correlation, his performance of his means action may already directly further his own goal, but at least the other's action performance will further his goal.⁵ Because of the latter, inbuilt feature, each cooperative means action will at least in that indirect sense further one's own goal. In general, if mutual cooperation can be expected to further the participants' present goals – or some of their other goals – better (or at least not worse) than what their acting alone could involve or bring about (cf. rational exchange where the participants reward each other), then cooperation is rationally undertaken.

Above the agents perform the token actions x_1 and x_2 . We may now call their collection x_1, x_2 , or in the n-person case x_1, \dots, x_n , an *individually inten-*

⁵ That the commonality of goals, interests, and preferences is central for cooperation has been argued for in Tuomela (2000 Chapt. 9), where a technical index of correlation, *corr*, is used to measure the correlation or correspondence in question (*corr* is a suitably normed Pearson coefficient). There the following thesis is defended: *Commonality Thesis*: Other things being equal, the more commonality of interest (as measured by the index *corr* and preferences) there is in a situation, the more likely cooperation is to be initiated and carried out successfully and – speaking of reward-based cooperative situations – to give the expected rewards, understanding this to mean rewards from acting together (relative to not acting so). That cooperation is successful in the sense of the Commonality Thesis has to do with improvements related to

- (a) the selection of a collective goal and the means of reaching it,
- (b) the stability of the commitment to the collective goal-directed action,
- (c) the possibilities of helping, and
- (d) flexibility concerning the change of a collective goal when it is called for.

In relation to (CIM) we can speak of its clauses [1] and [2] constituting a cooperative goal (in the thin sense discussed in the text) and this goal qualifies as the kind of goal spoken about above

tional cooperative collective action or, using our technical terminology, an *intentional cooperative collective action in the I-mode*.

Let us still emphasize that acts constituting cooperation are necessarily intentional. In terms of our analysis we can say that the intentionality of cooperation depends on an intention to cooperate, possibly under the participants' own descriptions of cooperation involving the main elements of (*CIM*). This intention thus basically has the content to perform means actions of the kind specified by clause [2], given the information in clause [1]. The intention to cooperate accordingly does not entail that the participants need to have a separate goal to cooperate over and above their intended goals and intended actions as specified in [1] and [2] of (*CIM*), rather it can be seen as an intention-in-cooperation when those two clauses are fulfilled. It is not required that the intention be reflected by the agent and represented in his mind. Note that cooperation of course does not guarantee that G_1 and G_2 actually are achieved.

One can also speak of a third-person view of cooperation. This we would call *quasi-cooperation*: the participants act as if they were cooperating with each other although they were not, they perhaps did not believe that they were cooperating and did not intend to achieve anything by acting with other participants. This kind of quasi-cooperation could also occur when the result is intended to be achieved by one person but in fact is achieved by many persons.

In I-mode cooperation the agents have privately intended proximate or further goals. Such a goal entails a private commitment to act accordingly. However, it must be noticed that the notion of having the further goal is aspectual, a goal is had under certain descriptions and here the goal, say making one's town beautiful, is not had under the description that it is the group's goal. Below in the we-mode cases, when the further goal is collectively accepted to be the group's goal, it is at least a part of the group's ethos (or at least entailed or implicated by it). In the I-mode case the further goal, when shared, is a counterpart to the group's ethos in the we-mode case and thus it could be called an "I-mode group ethos".

When people act together in the I-mode (in the sense of (*CIM*)) they do not act as a group, which they can be said to do when they act in the we-mode. In the I-mode case they act collectively in an "aggregative" but dependent sense. In the I-mode case a person's private further goal is achieved only when he has achieved it by his action. In contrast, in the we-mode case the Collectivity Condition holds (cf., R. Tuomela 2002a, Chapt. 2). Applied to goals, it says that, necessarily (due to group acceptance) when the goal is achieved for one person it is achieved for all, viz. the goal state token satisfies everyone's goal.

Our (*CIM*) allows for I-mode cooperative action in a group context. We get such a context in a rudimentary sense by assuming that $G_1 = G_2$ and they form the ethos of the simple I-mode group that the participants here can be taken to form. In more interesting cases there is a pre-existing group that the goals G_1 and G_2 serve (e.g., the ethos could be the goal of cleaning up the Gulf of Finland, and the participants' goals would be the personalized forms of this goal, A_i 's goal being to do what he can to help the ethos goal being achieved). The end result of the achievement of the goal is attributable to all

actors; and in this sense they can be taken to be collectively responsible in a causal sense for the end result.

The degree of social connectedness intentionality in cooperation may vary as well as the social “glue” between the parties. Mutual belief of cooperation makes for stronger intentionality in the cooperative action and a stronger social “glue” between the participants than mere shared belief. A promise-based I-mode collective cooperative action is the strongest case with regard to the intentionality of the cooperation and the social ties between the actors.

According to [3] of (*CIM*) the agents are required to have a mutual belief concerning the relevant pieces of information. Here is an argument for clauses [3] and [4]: suppose that the participants’ goals and actions are dependent in the sense of clause [2], be the dependence antecedent or voluntarily created on the spot. Then the participants will rationally need the other’s performance at least for an optimal result (viz. goal achievement) for themselves. This rationally requires that each agent A_i , $i = 1, 2$, believes (and, in addition, as will be seen in Sect. 4, “predictively” trusts) that the other one will adjust his goal and perform his means action as required by clause [2]. In addition, A_1 must of course believe that A_2 believes that A_1 will similarly participate, for otherwise A_2 would not have good reason to participate; and similarly for A_2 . Furthermore, they must lack the belief that the negation of the relevant higher-order belief contents are true. (See Tuomela 1995, Chapt. 1, and Tuomela and Miller 1988, for this kind of analysis and justification for clause [3].) To justify clause [4], we note that the resulting mutual belief cannot be idle: it must serve as the participants’ (partial) reason for participation.

However, there can be minimally rational cooperation even without clause [4]. Clause [3] serves to make cooperation to some extent social and rational, because it gives a reason to say that everyone believes that they are cooperating: “All of us are in it – my further goal is promoted also by the others’ promoting their further goals”. The cooperation thus is socially intentional as there is social awareness about the participants’ goals and actions: “I act to achieve my proximate goal of doing all I can to further my aim to make the town more beautiful and I believe everyone has the same proximate goal and the same or shared ethos and is acting in a parallel way and that there is mutual belief about it”. The actions form a cooperative collective action, as the agents share a further goal which they promote by their intentional actions. This is still an I-mode collective action, which is not collectively intentional. Here the participants have the shared I-mode intended we-goal as their partial reason for acting (in terminology of R. Tuomela 2002a). Briefly, a we-goal is one that the agent has in part because (he believes) that the others in the group have and are mutually believed to have.

With clause [4], cooperation is privately and individually intentional for all the agents, and the others’ acting is a (mutually believed) partial reason for one’s own acting. Here is an example: I intend to decorate my house partly in order to make the town beautiful, and a partial reason for this intention is that I believe that the others intend similarly (in a reason-based sense) and that there is a mutual belief about all this. The social aspects of cooperation here contain mutual belief as a social reason, and it is a case involving a reason-based intended we-goal in the I-mode.

In promise-based I-mode cooperation the participants are only privately committed to take part. The private commitment is a social normative commitment due to the mutual promises or acceptances. Each person has normatively committed himself not only to himself, but to the others to do what he has promised to. If a participant wants to pull out from the cooperation he is accountable to the others. However, to use an analogy, he is the “captain” of his boat and he views the situation from his own point of view and not necessarily from the point of view of the group’s interest. He is steering his boat in the same direction as the others as long as it seems worthwhile for himself, but he is not a co-captain sharing authority with others on the same ship in view of the group’s interest. Such sharing of the authority over a group’s action requires collective acceptance and entailed collective commitment, as in the we-mode case to be discussed below; and without the collective commitment to a joint goal there is no group-perspective in a full sense and no acting as a group member in a full-blown sense (viz. in the standard sense of Sect. 2).

Nevertheless, in the present I-mode case the actors may be taken to think: “I will act with the others to promote my further goal which they share, in part because I promised to do so and because I believe that they will act and think in the corresponding way, having promised the corresponding thing, and all this is mutually believed among these persons. Here the intentionality (in the social connectedness sense) of the cooperation is as strong as can be in the I-mode case, as their I-mode intentions are tied together by their promises.

In the present case of rational cooperation, the social aspects in the present case contain mutual promise and mutual knowledge about others’ intentions in addition to the ones in previous categories. Here we have a strong I-mode we-attitude and action on it.

To end this section, we consider briefly how (*CIM*) functions, and we start by giving a simple example to illustrate it in the case of a *coordination* situation. Agents A_1 and A_2 have as their shared private goal to meet (thinly conceived: to arrive at the same place). Let the equally good means actions in the case of both agents be going to the station (s_1 , s_2 , respectively) and going to the church (c_1 , c_2). Their actions leading them to meet (viz. the pairs s_1 , s_2 and c_1 , c_2) achieve coordination and lead to the satisfaction of their goals. The two other action pairs do not satisfy their goals. There is no conflict in this situation. Both agents are assumed to act individually (“privately”) rationally to successfully achieve their goals. This entails that they must have coordinated successfully (viz. selected one of the action pairs leading to coordination). Here the participants are dependent on each other in a regress-creating way, as emphasized in the literature (see e.g., Sugden 1993 and Tuomela 2002b). Here the rational row player will reason with respect to the column player (“you”) as follows concerning an action alternative X: I will do X if (given that) you will do X, but you will do X if I will do X, but I will do X if you will do X, but you ... *ad infinitum*. There are both unilateral and multilateral ways out and these ways need only be I-mode ways (in the sense of (*CIM*); cf., Tuomela 2002b).

Note that we have the same regress also when there is difference in the positive payoffs and we have what has been called a “Hi-Lo” situation. Here the coordination problem can be solved in the I-mode e.g. if one participant

commits himself to choose Hi and signals this to the other one. Then it is rational for the latter agent to choose Hi as well, and this fact makes the former agent's action rational. Another way to go would be to assume that it is mutual knowledge among the agents that each unconditionally acts on the basis of dominance. This is, however, a contingent assumption the rationality of which can be questioned.

Conflict can be introduced into a coordination situation by assuming that the agents have different preferences concerning where to meet, although they still prefer to meet rather than not. Here the conflict is not disturbingly big, and this modified coordination case (a "Battle of the Sexes" situation) qualifies as I-mode cooperation in the sense explicated by (*CIM*). In general (with some exceptions) a Prisoner's Dilemma cannot be solved cooperatively in the single-shot case in the sense of (*CIM*) without changing the game. However, a centipede game (an extended Prisoner's dilemma) offers a partial exception to this, for "long term" rationality in conjunction with pro-group I-mode action can lead to a cooperative solution. Players thus can get to the last choice node with long term rationality (in a sense involving assurance of the other's cooperation), but the last choice requires taking pro-group action and here acting on an I-mode (or possibly we-mode) utility, against one's initial I-mode utilities. As shown in Tuomela (2000, Chapt. 11), Hume's farmer's dilemma (a centipede situation and an extended PD) can thus be resolved by acting for an I-mode group reason (cf., "I help you because you are now one of us, no matter if we never meet again"). Note, however, that a centipede involves several moves but only one play of the game, so it is not fully a one-shot case.

In general, however, group reasons are likely to be we-mode reasons, and, as has been argued at length in Tuomela (2000), acting for collective or group reasons (e.g. towards a shared collective goal) may lead to even better results (e.g., instrumental gains related to costs, use of energy, time, etc.).

3.2 We-mode cooperative collective action

We argue that we-mode cooperative collective action is equivalent to acting together (jointly) towards a shared intended collective goal. This thesis is the central thesis in Tuomela's theory of full-blown, viz. we-mode, cooperation (see Tuomela 2000). The central idea here is to see that acting together (or joint action) indeed is in part constituted as cooperative activity. The shared collective goal in joint action serves to make the participants' part actions cooperatively connected. While there can be some conflict between the parts (cf. a game of tennis), the basic setup is that of cooperative harmony.

Intended shared collective goals are goals satisfying the Collectivity Condition; and, as goals can be regarded as contents of intentions (more precisely, "aim-intentions", in the sense of Conte and Castelfranchi 1995), we are actually just saying that we-mode cooperation amounts to acting together on the basis of a relevant joint intention. Consider thus this stylized summary of a simple two-person case, which abstracts from the underlying

reasons for joint intentions (goals), which reasons are taken to be positively correlated preferences (in a specific sense) specified in Tuomela (2000) (cf., (*WM*) of Subsect. 2.2, which is easily seen to be satisfied below):

(*CWM*) A_1 and A_2 (*intentionally*) cooperate with each other in the *we-mode* in bringing about goal G
if and only if

- [1] G is a collective goal type, viz., an “achievement-whole” the achievement of which can be divided – either *ex ante actu* or *ex post actu* – into A_1 ’s and A_2 ’s parts;
- [2] A_1 and A_2 jointly intend to achieve G by acting together and achieve G together in accordance with and partly because this joint intention of theirs to achieve G together.

To go from the above conceptually minimal notion to a notion of *rational* we-mode cooperation, we add two clauses saying that the information in clauses [1] and [2] must be mutually believed (clause [3]) such that the mutual belief is not idle (clause [4]):

- [3] A_1 and A_2 rationally mutually believe that [1], [2];
- [4] [2] in part because of [3].

Clauses [3] and [4] can be justified as in the I-mode case, *mutatis mutandis*, and we will not here comment on this matter. (Notice, however, that in addition to predictive trust in the present case “genuine” trust will be involved in the case of normally rational agents, as will be argued in Sect. 4.)

In (*CWM*) the collective intentionality of the action is either intersubjective (when based on mutual belief) or “objective” (when based on agreement making and mutual knowledge). The concept of acting together occurs in the analysans and should be possessed at least in a rudimentary sense by the participants. (In Tuomela 2000, the fullest kind of we-mode cooperation will also assume that the parties truly *willingly* perform their parts; here we will not consider this additional element of cooperation.)

We are here dealing with an at least intersubjectively shared, intended joint goal. The joint intention to achieve G is taken to amount to shared mutually known we-intention. The we-intention to achieve G together consists of a participant’s intention to participate in achieving G together with the others (so that the joint intention expression “We will achieve G together as a group” is correctly applicable) and of his beliefs that the right “joint action opportunities” (involving e.g. that the others indeed will participate) and that there is mutual belief about this (see Tuomela 1984; Tuomela 1995; Tuomela 2000; R. Tuomela 2003b; Sugden 1993 about this).⁶

⁶ Conte and Castelfranchi (1995) presents a goal-based account of cooperation based on dependence and a common (or “identical”) goal. This account bears some resemblance to the present account. An identical goal in their sense is roughly a we-mode goal in our sense. They require strong dependence between the agents and their actions. For a detailed discussion of this account see Tuomela (2000). Bratman’s account (in Bratman 1992) requires mutual responsiveness. For a critical discussion see Tuomela (2000) as well

(*CWM*) relies on the notion of jointly intentional joint action. This is collective action based on the participants' joint intention. Saying this leaves it open what a joint action simpliciter is. Basically it is action in which the participants have shares to perform such that their performances of them will create the "achievement" or fulfillment of a criterion definitory of the action (cf. carrying a table, singing a song together, or playing a game of tennis together). We will not here enter a deeper discussion of the notion of joint action except for the reminder that a joint action can involve some conflict as long as there is a suitable "joint action bottom" (see Tuomela 2000, Chapt. 3, and Tuomela 2004b, for recent discussion).

Given (*CWM*), the participants form a simple we-mode group already in view of their joint intention. As earlier, we assume that this group is both externally and internally autonomous. The participants need not be in fully symmetric position, there might be e.g., power relations involved specifying a task division. Still (*CWM*) applies, as the asymmetries are due to the participants' collective acceptance (recall our assumption of external and internal autonomy). It can be pointed out that in some rudimentary cases there can be certain "looseness" here. First, clause [2] entails that each participant has a we-intention to participate, and this can be understood liberally so as to allow that the participants need only have a we-intention-in-action to participate in the collective achievement of *G*. Thus the joint or collective *telos* in their part performances is to participate in acting together relative to *G*. Here *G* is the participants' intended *telos*, although not all the participants in many-person cooperation have this end as long as the cooperative action functionally works, so to speak. Of course, goal *G* need not be achieved for there to be cooperation. (The Collectivity Condition must also be understood liberally enough to become satisfied in these kinds of more general cases.)

In the present case of collectively intentional collective cooperative action the group members in effect have a collective intention rephrasable as "We will cooperate to achieve goal *G*". This contrasts with the I-mode cases in which the intentions are private and have the form "I intend to contribute towards (further) goal *G*" or, in the case of proximate goals, "I intend to achieve (proximal) goal *G**". We will assume that the group aims to promote its ethos and maintain it (as long as that is found to be rational or feasible). Goal *G* thus will typically be the content of the ethos of the group, and we will here speak of promoting the ethos on a general level. Taking the group, *g*, in question to be capable of action qua a group, we assume that *g* intentionally sees to it or brings it about that *G* is satisfied or promoted, or at least tries so to act. We can also, alternatively and equivalently, say that the action here is describable as *g*'s acting so as to promote its ethos. This is the case when its members jointly (try to) see to it or bring about that the ethos is promoted – based on their joint intention "We together will see to it that our goal *G* is promoted". When they together see to it ("stitt") that *G* is promoted, every group member can be said to have a part or "slice" in this joint stitt'ing, and this part is in principle irreducible to actions described without reference to the joint stitt'ing. We can now see that the above account (*CWM*) will be satisfied, given our assumptions. Clause [4] requires some rationality, though, and to cover all possible cases, this clause need not be required for

cases involving irrationality or, on the other hand, unconditional cooperators.

Even if on the group level, so to speak, the members are jointly seeing to it that G, they may act separately or jointly to achieve this goal and use whatever “tools” (e.g. hiring agents to do something) that are believed to be useful. Here the members’ part actions simply are actions qua a group member in the standard sense, and by these actions they *represent* the group. Representation here is based on (justified) *attribution* – their actions are attributable to the group in part because they are autonomously acting full-fledged members of the group. In internally autonomous groups the members can be regarded as authorized representatives of the group as long as they act as group members in the standard sense, and the members either all are operatives in equal positions or they have internally appointed some operatives for various tasks, resulting in a (normatively) structured group. The totality of the members’ part actions, based on their jointly stit’ing G, when successful, collectively taken constitutes the group’s ethos-satisfying and ethos-maintaining action. This is a we-mode cooperative collective (or joint) action.

There is shared joint (or, here equivalently, collective) intention towards the ethos in the present case, or towards another joint goal in other cases, and hence the participants are collectively committed to it (in our example case bringing about that the town is beautiful).⁷ In this kind of case each agent’s intention is to act so as to perform his part of their bringing about the joint goal as his part of it, or to do his part of promoting the ethos, and to achieve the goal (or promote the ethos) together with the others. The collective commitment entailed by the collective intention has the content that the members, as group members, collectively promote the shared collective goal (or the group’s ethos). In the case of rational agents, collective commitment presupposes that the persons believe or know that there are others who are analogously committed and that there is at least mutual belief about all this. We have argued that mutually believed collective commitment enables a group to act as a group and thus to perform we-mode collective cooperative actions with a full group-perspective (see Tuomela and Tuomela 2003, Sect. VII).⁸ Thus, in principle for persons acting out of the group-perspective there will be no collective action dilemmas, for there can be no

⁷ That the ethos can in general be seen as a goal is due to the fact that for the group members it has the world-to-mind direction of fit when its satisfaction is concerned. So, for instance, group values can be taken be collectively accepted in this precise sense

⁸ Margaret Gilbert has in her 1989 book and later works similarly emphasized the role of “joint commitment” in contexts of joint action and related contexts. Here notion of joint commitment, a we-mode notion in our sense, entails performance obligations to the joint action and to other participants. This contrasts with our we-mode notion of collective commitment, which as such is intention-relative and entails only a rather trivial “instrumental” or “technical” “ought”. Thus our treatment covers more ground than hers. In his new paper (2003), Christopher McMahon compares Gilbert’s notion of collective commitment, which involves that a decision to rescind must be a joint one, with what we would call I-mode collective commitment or “aggregated private” commitment. Such I-mode commitment is weaker in that a single participant can himself make the judgment that it is feasible for him to leave the joint action or that it would be reasonable for the group of participants to leave the project

conflict between individual and collective rationality in the group: the group rationally prescribes mutual cooperation for its members.

However, the reader should be reminded that in the case of collective action dilemmas, it may nevertheless be hard in actual life to bring about collective commitment toward the ethos, as a free-rider problem affects most such dilemmas (e.g., PD). What is required for success here is that the participants' preferences are shifted towards the ethos (and consequently toward others' preferences) so that in effect the dilemma is transformed into a milder dilemma without a free-rider effect. In the case of the PD, one typical such game is the Assurance Game (where the joint outcome of mutual cooperation is now preferred to the joint outcome where one defects while the others cooperate, reversing the corresponding preferences in the PD). We will not here discuss the ways in which this kind of transformation can plausibly and rationally come about (see e.g., Tuomela 2000, Chapt. 12, and Tuomela 2004a).

The full group-perspective is partly described by the slogans: "We will do it" and "All for one and one for all". The group's interest is in focus and not the private interest of the members, although they may further their own interest by furthering the group's interest. The success of the group's acting as a group is more important than the success of a member's acting. When people see matters from the group's point of view, private conflicting interests have to be laid aside. In I-mode cooperation people may come and go as they themselves see fit, at least if there is no agreement to cooperate. And when there is one, they may yet decide to withdraw when the project becomes costly and pay the fines for breaking a promise. In contrast, in we-mode cooperation the persons have "entered the same ship", trying to steer it to the chosen port as "co-captains" with shared authority. Shared authority presupposes that one person is not allowed to act as he chooses. His actions are open for all to criticize and he has to be prepared to receive assistance. This group-perspective is involved in the collective commitment to the group ethos, which necessarily involves commitment to acting for the group ("forgroupness", cf., R. Tuomela 2002a, Chapt. 5). From it follows that when the goal is satisfied for one person it is satisfied for all with necessity based on group acceptance, because the goal is the group's goal. The collective commitment involves social commitment due to the fact that the committed persons are co-authors of the action. Thus they are committed, not only to themselves, but committed to each other to performing their actions promoting the group's ethos. People ought to act as proper group members if they are to be consistent with their collective commitment.⁹ The social commitment component of collective commitment brings in a social "ought" (directed towards the fellow participants) to enrich the basic "ought" in collective commitment (directed towards the ethos and derivatively one's ethos-promoting actions).

In the case of mutual belief-based we-mode cooperation there can be collective intention and commitment, as from the agents' point of view

⁹ According to Tuomela and Tuomela (2003), there is an "ought" to maintain the group's ethos by proper actions as long as it is collectively accepted (whatever that in each case involves). Still, the ethos may of course be changed by the group

“things seem right” for shared we-thinking (even if objectively the beliefs might turn out to be false and the agent might find out that this is the case). As a consequence, what a person does within the relevant framework is not his private business, as the authority for the joint action is shared. His actions are part of a joint venture and open for the others’ scrutiny. He should be ready to receive help and criticism in service of the joint outcome. This transparency is strictly instrumentally grounded, but gets a social dimension from the fact that collective commitment involves social commitment – as discussed earlier. Note, however, that people may be mistaken about which persons in fact are among the properly committed group members. This may cause difficulties for achieving a goal where every member’s participation is needed. Here, there is no objective justification for sanctions in the case of defection. Social aspects in this category contain all the previous aspects mentioned in the context of (*CIM*), except the promise, and with the addition of collective commitment to a group’s goal or ethos.

The mutual belief-based case contrasts with the case of collective social cooperative action based on an agreement or a normative plan. Here the agreement is to promote the group’s ethos or to promote the group’s goal as a group, each person acting as a group member. There is mutual knowledge of the expressed commitment. People ought to be committed and can be expected to be disposed to be committed. The focus is explicitly moved from participant to group: “We will do it!” The group’s goal could be such that it cannot be achieved alone or by I-mode promise-based interaction where each person is “captain of his own boat”. The “glue” of strong private and social commitment is not always sufficient. A full group-perspective is needed e.g., for situations where a participant as the captain of his own boat would change the course of his boat, if he viewed the matter from his private point of view, instead of the point of view of the group’s interest. As a co-captain in the group-boat his central interest is to participate in achieving the group’s goal or promoting the group’s ethos.

While collective commitment to a joint goal or the group’s ethos entails at least a “technical ought” to do one’s share, in the agreement-based case there is also a “moral ought” to refrain from harming others by defection. As to the social commitment involved, the agreement makes it two-sided and objective. In contrast to the mutual belief-based case, there is here also an objective basis for justification of sanctions in the case of defection. Especially in agreement-based cases, a member should be released from the agreement by the others, but also in cases based on mutual belief where the cooperative action is going on, he should not leave without consulting the others. The agreement-based case is the strongest category for cooperative action and could be said to involve a strong we-mode we-attitude. The social aspects involved are as in the previous case with the addition of agreement and mutual knowledge.

This concludes our treatment of intentional cooperative collective action. As seen, all the cases considered in this section, viz. all I-mode and all we-mode cases and, can be regarded as collective actions based on a shared we-attitude. In this sense cooperative collective action is conformative action, which should not be a surprise. However, there can be conflicts in all these cases. Even in the strongest we-mode case of agreement-based joint action

there can be conflict. For instance, some group members may compete (cf., rivaling salesmen in a business organization) and be in partial conflict. But their positions in the organization are still consistent or, if competition is normatively built into the positions, at least there is much shared common ground ensuing from the shared ethos (basic goals, etc.) of the organization to which they are collectively committed.

To end this subsection, we show that acting as a group member in the core sense and cooperative collective action are equivalent on the basis of our analyses (*CIM*), (*CWM*) and the definitions (a) and (b) of Sect. 2 defining acting as a member in the core sense. The thesis is this:

(*COOP*) A group member acts as a group member (in the core sense) if and only if he intentionally cooperates with respect to the group's ethos with the group members.

We will discuss both the I-mode version and the we-mode version of (*COOP*). The we-mode version is trivial, although not completely circular, because the concepts of acting as a group member and cooperation with respect to the ethos of the group are closely related.

We-mode Thesis. A group member acts as a group member in the standard sense if and only if he cooperates in the we-mode (in the sense of (*CWM*)) with respect to the ethos with the group members (presupposing that also they cooperate).

I-mode Thesis. A group member acts as a group member in the weak sense if and only if he cooperates in the I-mode (in the sense of (*CIM*)) with respect to the ethos with the group members (given that they cooperate).

As to the we-mode thesis, the proof goes roughly as follows from left to right. Acting as group member in the standard sense entails that the acting members are collectively committed to promoting the ethos. We-mode cooperation here is cooperating together toward a collectively accepted goal (here: the ethos). Cooperating together here is we-mode joint or collective cooperative action in the sense of our (*CWM*). But this we have from acting as a group member, for the group members are supposed to do their share of the group members' collective action of promoting the ethos with collective commitment in conditions of mutual belief. Collective commitment to the ethos in the present context entails that the ethos is the members' we-mode collective goal, thus jointly intended goal. (*CWM*) thus applies to the present case.

As to the converse entailment, we-mode cooperation toward the ethos trivially gives us collective acceptance of the ethos, because the group's jointly intended goal required by (*CWM*) is plainly assumed to be the ethos and because joint intention entails collective commitment (and, conversely). Thus the collective action considered will involve these elements and the account (b) of acting as a group member in the standard sense is seen to hold true.

As to the I-mode thesis, we consider the implication from left to right by assuming action towards the ethos with private commitment to the ethos. Why do we now have I-mode cooperation? The simple proof is that we have parallel action towards the same goal, and this is minimally cooperative action if there is some mutual adjustment of action towards the others' goal. Acting as a group member in sense (a) entails that the agents are acting

towards the same goal in terms of the actions in classes [1]–[4] accepted by the group (here only in the I-mode). Action toward the ethos must by the nature of these actions be consistent and based on this “harmonious shared bottom”. The required adjustment towards the others’ activities will be there to some extent at least because the actions in [1]–[4] by their characterization promote the ethos, which here can be regarded as a privately shared goal. A member’s so acting towards the shared ethos will thus make its satisfaction likelier both for him and the others. Thus at least minimal cooperation in the sense of (*CIM*) is obtained.

To get the entailment from right to left we apply (*CIM*) to the ethos and consider the group members’ adjusting their actions toward the others in the sense of clause [2] of (*CIM*), assuming private commitment towards the shared ethos. Then the definition (a) of acting as a group member immediately applies and we have the entailment from right to left.

4 Cooperation and trust

4.1 Trust as a precondition and requirement for cooperation

Our first thesis in this section is that, at least on conceptual and rational grounds, *some kind of trust is needed for rationally initiating and maintaining intentional (I-mode or we-mode) cooperative action* between conditional cooperators. Rational people do not enter into cooperative action with partners that are likely to refrain from doing their share. A central problem is that a participant might not know that the others will cooperate. Systems of control and sanctions are costly but yet they do not always have a deterring effect on free riders. There may be loopholes and even when cheaters get to pay for their breaches there will be lost resources when the cooperative action is harmed. The existence of codes of honor is helpful, but only if it is known that your partners adhere to them. An inner commitment to fair play may be based on various motives, one of them being the (collective and individual) advantage resulting from cooperation. However, we cannot be sure about the inner commitment of others nor of their other motives, e.g., of how important the gain of cooperation is for another person. When knowledge is lacking, trust is needed for cooperation (as also e.g., the empirical research in Kollock 1994 indicates).

In the literature, trust is spoken of in different senses. Thus, “genuine trust”, to be discussed below, is typically assumed to be restricted to close relationships while “predictive trust” or “predictive reliance” belong to the framework of interaction between people in general.

According to M. Tuomela (see M. Tuomela 2002, 2003b, and especially M. Tuomela 2003a), genuine trust is rational social normative trust (*RSNTR*) due to a relationship of mutual respect. The account of rational social normative trust (*RSNTR*) for primary trust relations is presented below, starting with the account of the context for rational trust (*Y*):

The following conditions are necessary and jointly sufficient for *a context (Y)*, “*the trust context*”, in which A can rationally trust B that he will perform a specific action X:

- [a] A wants B to perform X, an action that A does not intend to perform himself. A thus believes that he is concerned with B's performing X, and believes that he is in that sense dependent on B's action;
- [b] A has a belief about B's belief of the dependence of A; and
- [c] A believes that his dependence is "relevant" to B. That the dependence is relevant to B means that B is free to allow A's dependence to play a role in his decision-making, and that B takes A's dependence into account in his actions.

A's beliefs in [a]–[c] may be dispositions to have such beliefs and they are based on (at least) subjectively rational reasons (or dispositions to have such).

(*RSNTR*) For a person A to have *rational social normative trust* in another person B that he will perform a specific action X in a situation in which context Y is present, the following necessary and jointly sufficient conditions must hold.

- [1] Intention condition: A expects (or has the disposition to expect) that B intends to perform X. The expectation is based on (at least) subjectively rational reasons or dispositions to have such.
- [2] Ability condition: A expects (or has the disposition to expect) that it is possible for B to perform X (in that the internal and external opportunities for B's action obtain). The expectation is based on (at least) subjectively rational reasons or dispositions to have such.
- [3] Goodwill condition: A expects (or has the disposition to expect) that B will intentionally act with goodwill, including goodwill towards A when performing X. A believes (or has the disposition to believe) that, due to their relationship of mutual respect, he is entitled to expect this of B, on social normative, or (quasi-) moral, grounds, and that B acts at least in part because A has this right. The expectations (or dispositions) are based on (at least) subjectively rational reasons or dispositions to have such. The reasons for A's expectation in [1] may be (partly) the same as here.
- [4] On the grounds of [1], [2], and [3], A has a "positive" feeling. This feeling is or at least includes the feeling of being comfortable about his dependence on B for performing X.
- [5] On the grounds of [1], [2], [3], and [4], A has an accepting attitude vis-à-vis his position of dependence on B for performing X. This attitude is non-intentionally acquired.

The account of predictive "trust" (*RPTR*) is the same as above except that it lacks the belief about the right based on the relationship of mutual respect in its corresponding clause [3].¹⁰ Predictive reliance lacks context (Y)

¹⁰ See Castelfranchi and Falcone (2001) on a rich and detailed account of trust that resembles M. Tuomela's notion of predictive trust (viz. trust as "depending on" features of a person or a situation). In Sect. 2.3, they analyze trust as an attitude, "Core trust" that consists of beliefs, and trusting as a broader notion including "deciding" to rely on another person. Their first notion is weaker and their second notion is stronger than the notion of predictive trust discussed above

and the goodwill expectation concerns a third person or no person at all. In the latter case “the action performed with goodwill” means that the actor is highly committed to what he purports to do. Predictive trust and predictive reliance may also involve a normative expectation, but this kind of trust and reliance amounts to merely depending on the features of a person and/or the situation. Predictive “trust” is used as a term involving a person’s belief in another person’s intentional gratifying action, while predictive reliance is used when the gratification of the relier is believed to be merely a side effect of the other’s action. In the latter case people may intentionally enter a rudimentary cooperative action, believing that the others’ actions will, as a side effect, produce what the former need for achieving their aims. They adjust their actions to the others’ actions only to facilitate their own actions but (possibly) not the others’ actions. This rudimentary cooperative action is not fully intentional in a social sense or intentional as a cooperative action even in the sense of (*CIM*), as the parties merely are making use of each other’s actions as anticipated pieces in their own puzzle works. Here they merely relate to each other as rational, strategically acting “objects”.

In contrast, in cases of predictive trust people expect that others will perform their actions with the intention of doing what is called for, taking into account what others expect of them (making clause [2] of (*CIM*) satisfied). Here people relate to each other as intentionally cooperating subjects. In predictive trust a person’s motives for cooperation are evaluated as well as the effect of systems of control and sanctions. Evaluation is always less than perfect as we cannot really read the minds of others. This kind of depending on features of the person or the situation is, despite feelings of security, acknowledged risk-taking from a first- and a third person-point of view and does not deserve to be called genuine trust. In social normative or genuine trust, which will be presented in the next subsection, the trustor does not conceive of any risks, although the trustee has the freedom to disappoint the trustor. This is because genuine trust is not the result of evaluation, but is rather due to a relationship of mutual respect, where trust is a constitutive aspect. The evaluation is one step “below” or “behind” – connected to the relationship viewed as one of mutual respect.

The thesis that – on conceptual and rational grounds – some kind of trust is needed for rationally initiating and maintaining an intentional cooperative action between conditional cooperators gets support from the following reasons. The central element in the analysis of a person’s (A) rational predictive trust in another person (B) vis-à-vis B’s performance of an action is A’s belief that B will intentionally perform the action with goodwill towards A. B’s acting with goodwill toward A should be understood as B’s acting so as to accommodate A in his want concerning B’s performance of the action (this connects clause [3] of (*RSNTR*) and the corresponding clause in (*RPTR*) to clause [2] of (*CIM*)). A thus believes that B will act intentionally for a social reason. That this type of a belief must be present in rational cooperation, both in the I-mode and the we-mode case, was in fact already shown on the basis of the argument given for clauses [3] and [4] of (*CIM*) in Sect. 3. (The mutual belief in clause [3] entails the belief in question.) One may analogously comment on clauses [3] and [4] of the analysis (*CWM*) of we-mode cooperation. However, we wish to emphasize that, as will be shown

in Subsect. 4.3 below, we-mode cooperation in normal rational cases involves genuine trust in the sense of (*RSNTR*) and our proof there relies on somewhat different considerations. Accordingly, in the present subsection we will concentrate on I-mode cooperation and on showing that it requires either predictive or genuine trust.

To continue the argument, it appears to be quite trivial to state that each rational party in cooperation has to believe that, in order to initiate the cooperative action and to go on cooperating, the other will intentionally do something that furthers the interest of his partner. However, we have claimed that the mere Goodwill belief of rational predictive trust (*RPTR*) is not sufficient, but that at least predictive trust is needed for rational cooperation. The analysis of rational predictive trust involves some other conditions, viz. the conditions of (*RSNTR*) with the exception of that part of clause [3] which concerns respect and rights, and the satisfaction of those conditions is also needed in the case of rational agents. The conditions of context (*Y*) and the first two belief clauses of the trust account have to be satisfied for reasons of rationality. The fifth condition has to be satisfied for making cooperation possible. If a rational person believes that his potential partner will do the wanted thing, but the believing person still does not have an accepting attitude concerning becoming dependent on his partner (clause [5]), he is not only inconsistent (nonrational) but he is clearly not ready for cooperation. Note that trust does not require action, but provides readiness for action. Thus clause [5] is needed for initiating and maintaining cooperation not only on rational but also on conceptual grounds and presumably on psychological grounds (or else these parties have to rely only on their own action). The satisfaction of the fourth condition is not needed for cooperation, but lacking a “positive feeling” about the future dependence on another person may well turn potential cooperators to seek other partners. Anyway, it is “rational” to react with positive feelings to becoming dependent when gratification is expected.

Predictive trust could be said merely to amount to “depending” on features of a person and the situation. Evaluation of features based on reasons is acknowledged risk-taking and all of us would, of course, prefer to avoid risks when cooperating with others. In genuine, viz. social normative, trust, to be discussed below, the trustor does not conceive of any risks, although there are risks seen from a third-person point of view. Genuine trust is stronger than an evaluation-based attitude. However, there is evaluation concerning the underlying relationship of mutual respect. This is typically unreflective evaluation as a part of a process where experiences are interpreted as signs of respect. Genuine trust also involves some less central, overtly predictive components related to the specific matter at hand (viz. context (*Y*) [b], (*Y*) [c], (*RSNTR*) [1] and (*RSNTR*) [2]). It requires much but gives a solid ground for cooperation.

4.2 Genuine trust in cooperation

According to the account (*RSNRT*) of social normative or genuine trust, the central elements of such trust are the following types of beliefs. A person *A*

genuinely trusts another person B concerning his performance of an action X, when A expects (or has the disposition to expect) that B will intentionally act with goodwill, including goodwill towards A, when performing X. Furthermore, A believes (or has the disposition to believe) that due to their relationship of mutual respect he is entitled to expect this of B, on social normative or (quasi-) moral grounds, and that B acts at least in part because A has this right. Thus, genuine trust is an attitude that arises and thrives in a relationship of mutual respect. Here each person respects the other's right to have what he thinks he should have himself if he were in the other's place. This makes the rights relationship-specific and sometimes even a matter of negotiation, although the rights typically are expected to be grounded in general or group-specific social or (quasi-) moral norms. Trusting outside a relationship of mutual respect is to depend on a person's features, such as, his love, kindness, moral character, or motives to gratify the trustor (viz. it is predictive trust).

To grant others the same rights as you think you should have in their place could be called a basic moral we-perspective or, more accurately, a basic "quasi-moral" we-perspective (as, strictly speaking, it does not in all cases guarantee a moral reason for action, although it expresses a kind of fairness idea). This is not the same as the we-perspective in the sense of a group perspective, which involves viewing matters from the point of view of the group, furthering the group's interest even when it does not coincide with the private interests of the members. However, our basic moral or quasi-moral we-perspective (roughly, a disposition to live by the Golden Rule) seems to be a wholesome and typical attitude in successful group life, and as will be shown later, it is indeed entailed by the group-perspective of the we-mode.

People may have predictive trust in persons who have a basic moral we-perspective concerning *all* his fellow human beings (viz. trust as depending on features of a person). Such Kantian (universally) moral persons who respect others' rights outside the local setting of secure mutuality will sometimes find themselves in the sucker's position, but probably more often they will be the ones who get cooperation started. A signal of cooperativeness or taking the first step will help those who hope for long-term cooperation.¹¹ However, cooperation based on predictive trust need not involve beliefs of the participants' moral character of any kind. The trust may e.g., rest on beliefs of the participants' prudential motives for taking part. In genuine trust the relationship of mutual respect (involving the basic moral we-perspective) is crucial. This kind of relationship can obtain between partners both in I-mode and we-mode cooperation. In the latter case the basic quasi-moral we-perspective is a structural element of the group-perspective (cf., Subsect. 4.3).

In a relationship of mutual respect the mutuality of the basic moral (or quasi-moral) we-perspective is part of the relationship and it is here that genuine trust can arise and flourish. In genuine or social normative trust

¹¹ In all simulation-based accounts of cooperation we know that some amount of "Kantian" cooperation must be present to solve the problem of cooperation and to prevent free riders. (cf., Gintis 2003, and Richerson and Boyd 2001.)

the trustor does not conceive of any risks, although it is a fact that the trustee is free to choose his course of action. The problem of whom to genuinely trust has moved one level down: how do we know when we have a relationship of mutual respect? It seems equally difficult to answer this question as the corresponding ones about friendship and love relationships. Only experience will show us the true nature of a relationship.

Likewise we would like to know how to acquire relationships of mutual respect. Mutual assurance about respect and desire for a long-term relationship is a declaration of commitment that still has to be brought into the challenges of practical life, but it is a good starting point for the relationship. Thus a relationship of mutual respect and genuine trust must be preceded by mutual risk-taking and predictive trust based on evaluation of the features of the person and the situation. Experiences of mutual (predictive) trustworthiness give reasons to believe that a relationship of mutual respect is being established.¹²

4.3 Successful cooperative action promotes, and we-mode acting entails mutual respect and rational trust

It has been argued in Subsect. 4.1 that at least predictive trust is rationally required for intentional I-mode and we-mode cooperation. The second, partly conceptual and partly empirical thesis of this section is that *successful rational intentional cooperation (as qualified below) promotes (subjectively) rational genuine trust*. This thesis is true contingently, due to the way the world is, but there are nevertheless conceptual connections between the antecedent and the consequent. Briefly put, we contingently suppose that rational successful cooperation will give a participant the experience of the other party having respected his right to rely on the latter's participation and contribution. When mutual, this experience, especially when repeated, provides the basis for rational participants to see their relationship as one of mutual respect in the sense required by (RSNTR). Having the appropriate experiential grounds it is a contingent but plausible assumption

¹² Acting with goodwill" and "having a relationship of mutual respect" are, of course, separate notions. In the framework of the account of rational social normative trust (RSNTR) the first notion refers to a way of acting (here in a gratifying way, accommodating the other person in his want), based on any kind of (self- or other-regarding) motive. The second notion involves the parties' view of their relationship as one where each respects the other's rights to be gratified by the other, based on their subjective experience of this being the case. According to the genuine trust account (RSNTR) the trustor believes, due to their relationship of mutual respect, that the trustee will intentionally act with goodwill towards him, performing the desired action, in part, due to his right to expect this. Thus, the trustee's respect of the trustor's right is expected to be cashed out as the trustee's intentional good-willed action, at least in part, for the reason of that right. This kind of acting with goodwill is only a subclass of the broad category of acting with goodwill. People may have predictive trust in others and expect them to intentionally act with goodwill towards them for a variety of reasons other than respect of their rights. Often there are no rights involved, or the expected good-willed action may go far beyond what any rights would warrant

that the mutual respect relationship is indeed formed.¹³ When this is the case, the participants have the central basis required for having rational genuine trust in each other regarding their contribution. There will be genuine trust, provided that the other, here rather unproblematic, conditions of (*RSNTR*) are satisfied. As will be argued, they are satisfied for normally rational agents in the context of rational cooperation. (Normally rational agents are assumed to have the knowledge, abilities, and mental dispositions that the common sense framework of agency assumes for its applicability.) As seen, a contingent condition for the above thesis about the promotion of genuine trust by cooperation is that the cooperation must involve some rights of the parties. In addition, in successful cooperation, “successful” here means not only unproblematic interaction but also that there is mutual respect of rights.¹⁴

What is conceptually true in the thesis is that, given rationality, experiences of having one’s rights respected in successful rational cooperation give a reason to believe that a relationship of mutual respect is being established, which relationship is a required basis for rational genuine trust. Such trust between the parties of such cooperation *vis-à-vis* their contributions will be promoted in a relationship of mutual respect, when the conditions of the trust account are satisfied (see the comments below). The above thesis can be strengthened in the case of *we-mode* cooperation, as will soon be shown.

We hypothesize that the participants’ relationship of mutual respect is historical in the sense that it cannot be acquired by any other method than by having experienced such mutual respect in a cooperative relationship. In special cases the relationship can perhaps yet be formed in a single round of cooperation, without a previous history of relevant encounters. Only when people mutually experience that their rights are respected and that they

¹³ There is some relevant empirical research on formation of trust, although to our knowledge genuine trust has not been directly studied in these investigations. Nevertheless, Kollock’s studies (in Kollock 1994) on the relationship between different exchange conditions and levels of trust indicate that increasing the frequency of positive interactions between partners indeed facilitates trust and does it even in conditions of certainty in which mutual trust is not so crucial. See also the relevant papers in Ostrom and Walker (2003), in many of which iterated game situations are studied in relation to formation of trust. We will not try to specify the exact conditions under which mutual experiences of respect indeed lead to a stable relationship of mutual trust but assume only that in general this will eventually happen under normal conditions of repeated cooperation, probably at different rates in the case of different persons and circumstances

¹⁴ The parties may be believed to respect each others’ rights for self-regarding or other-regarding reasons. As long as they have subjective experiences of having their rights respected, they get a chance to come to see their relationship as one of mutual respect. The first type of reasons would suffice for viewing the relationship as one of thin respect and make thin genuine trust possible, while the second type would suggest an understanding of the relationship as one of thick respect and make thick trust possible. Thin genuine trust could be seen as local in the sense that it is there as long as the reasons for maintaining the relationship are there, while thick trust is based on a relationship of a wider scope and durability. However there is no sharp line between the two, as self-regarding motives may give stable relationships of respect. It remains an empirical matter how many respectful cooperative actions typically are sufficient for convincing people that they are in a relationship of mutual respect regarding specific matters. At least, one can claim that each such action paves the road for a new attempt, and each successful experience adds up to an overall experience of having a relationship of mutual respect

respect the others' rights, they tend to conceive of their relationship as one of mutual respect. Their experience may be only illusory, but it forms the basis for their subjective view of the matter. If the view of the relationship is that the respect is one-sided, the person who experiences the respect cannot genuinely trust the other, as he understands that the asymmetric situation is not a stable and good relationship where genuine trust is a constitutive feature. He may instead have predictive trust based on evaluation of the likelihood of the other's persistence in his respectful action.

For taking the step between successful rational cooperation (with respect of rights) to promotion of genuine trust, we need to connect the experience-grounded belief of having a relationship of mutual respect to genuine trust. Every relationship of mutual respect, based on previous successful cooperation, does not automatically result in genuine trust vis-à-vis the contributions of the participants in a new potential case of cooperation. The other conditions of (*RSNTR*), however, arguably are satisfied for normally rational agents (viz. the participants have the beliefs and other attitudes that can be assumed for the applicability of our common-sense framework of agency) when indeed it is rationally justified for them to believe that the opportunities for cooperation exist. Given this, a certain amount and kind of experiences of mutual respect will be subjectively regarded by the participants to be sufficient for them to have a relationship of mutual respect, and given this empirical result, right-involving rational successful I-mode or we-mode cooperation will generate rational genuine trust among normally rational agents vis-à-vis their justifiably expected part-performances or contributions. (The major informative work in arriving at this result is of course done by our empirical hypothesis of respect generation.)

Viewing the various categories of cooperative action (Sect. 2), it seems that we-mode action is an action mode where mutual respect finds support in some structural features. This is because collective commitment, involved in the we-mode, also contains social commitment. Social commitment again involves that a participant has the right to expect that the others perform their parts and it also involves that this right is respected by the others, all this being mutually believed. The former claim is true because of a participant's obligation (be it only technical or quasi-moral) to perform his part. The latter claim is true when the participants indeed are socially committed: A person's actually binding himself to perform his part because of this obligation will yield action (part-performance) which amounts to respecting the other's right to expect that part performance.

I-mode cooperative action does not involve such structural features, but can of course involve mutual respect between the parties. Promise-based I-mode cooperation provides a structure for mutual respect by the promise. The parties of the cooperation give and receive a promise to the effect that their part will be performed. They are justified in expecting respect for their right to be accommodated. Mutual respect is intrinsic in the idea of mutual promising.

In cooperative action a free rider might get away without detection and some participants may withdraw their participation, but this should not stop people from entering cooperative projects. If people pass the test of evaluation and get predictive trust, they declare their commitment for long-term

cooperation in mutual respect, they should be allowed to become partners of cooperative projects on the basis of this “as if-trust”. Inside the framework of cooperation, sympathy and a commitment to moral values may pave the road to a true relationship of mutual respect and genuine trust. The experience of having had one’s rights respected and having respected the rights of the other in a cooperative action of any kind is crucial for acquiring the belief that a relationship of mutual respect is being established.

Those members who are collectively committed to the group’s ethos or joint goal ought to perform their share, not only for being consistent with their own intentions, but because the others have a right to expect that “every wheel under the wagon rolls”. Theoretically this follows from the assumption that the people are involved in a we-mode situation of cooperation. They are thus collectively committed to act together, and this involves social commitment towards the others (“being in the same wagon”). The we-mode entails the Collectivity Principle, which here amounts to collectivity concerning the respect of each other’s rights that are based on their social commitment to each other. That is, it is necessarily the case on non-contingent grounds that if a participant’s rights ought to be respected by the others then the same holds for every other participant. So we have here a group version of the Golden Rule principle. However, the right dealt with here need not be more than a mere technical right with a social aspect. The social we-mode structure of a group-agent does not entail *genuine* mutual respect (in which the respected rights are acknowledged as moral or social rights), but perhaps its technical counterpart: “I ought to respect your rights to expect my part performance as you ought to respect mine to expect your part performance, because acting jointly requires that each of us performs his part and for this we are accountable to each other”. The we-mode does not entail only the basic quasi-moral we-perspective of mutual respect, but most centrally a group-related we-perspective in which the participants are we-mode committed to further the group’s interest. The technical counterpart of respect may grow into genuine respect for members as members of the joint cooperative action. Here genuine respect stands for respecting the rights of others, acknowledging that the rights are moral or social. Note that thin genuine respect does not require other-regarding motives. A may have prudential reasons for respecting B’s moral or social right to have A’s contribution.

Some members may be mistakenly thought to be among those who are collectively committed to the group’s ethos. They are either free riders, individualists who are privately committed to the ethos, or dissidents. When groups rid themselves of problematic members the good structure of a collectively committed group may become a place where each member learns that fair play makes for rewarding cooperation.

The discussion of the mutual respect in connection to the mutual belief-based case is, of course, even more valid in the case where agreement is the objective sign of having entered a collective cooperative relationship. People have an objectively justified right to expect that the group perspective is adopted. “From now on we stand or fall together!” The participants have equal mutual rights and duties, which include monitoring others’ actions and helping when needed. They have voluntarily entered a framework, which functions as a relationship of mutual respect as group members, although

not as private persons. The basic quasi-moral we-perspective is true for collectively committed participants as group members (recall the mentioned connection between the we-mode, the entailed collective commitment, and the collectivity principle expressing a quasi-moral Golden Rule). Each person grants the others what he thinks he should be granted himself and what he demands of the others he demands of himself. A relationship of positional mutual respect is a natural part of a set-up where people are bound together by a full group-perspective. Thus the collectively committed members of a cooperative action ought to and are disposed to cultivate a relationship where genuine trust in members qua members can thrive. It could be called a thin genuine trust as long as the motives are instrumental.

We will claim as our third thesis that, *given some qualifications*, when some participants rationally cooperate in the *we-mode*, *where the members hence are sincerely collectively committed (involving, at least, an instrumental and social obligation towards the others to perform one's part, in a case lacking an agreement), and where they mutually believe that this is the case, their relationship is characterizable as one involving mutual respect and genuine trust, viz. (subjectively) rational social normative trust in the sense of (RSNTR), vis-à-vis their part-performances*. This thesis is true basically (but not completely) on rational and conceptual grounds. The qualifications needed are that the conditions of (RSNTR) other than the requirement of a relationship of mutual respect will be true. Those conditions will be true for normally rational agents in rational we-mode cooperation. In addition the contingent assumption is needed that cooperation indeed here guarantees the presence of an in-principle stable, non-fleeting social relationship of mutual respect. This assumption is normally true in a singular we-mode action, for there members normally conceive of their relationship as one of mutual respect, even without previous experiences of having had their rights respected.

The proof of the above thesis goes along the following line. The social commitment involved in collective commitment in we-mode cooperation involves that the participants not only expect or are disposed to expect that the others perform their parts, but that they have the right to such expectations and mutually know this (cf., Tuomela and Tuomela 2003). The social right here can be called “quasi-moral” as it concerns refraining from letting the others down. This phase of the argument thus has established part of the mutual respect assumption needed by (RSNTR). What is yet needed is to show that there is a non-fleeting social relationship of mutual respect. This was above said normally to be the case on contingent grounds.

According to the account of social normative trust, the central element of such genuine trust is a trustor's belief of the following kind. “Due to a relationship of mutual respect, he normatively expects of the others, on social or (quasi-) moral grounds, that they will act with goodwill towards him in those matters where he has a right to expect such action, and he also expects that they will act so, in part due to his right” (cf., clause [3] of (RSNTR)). When the motives become genuinely other-regarding this social normative trust turns into thick genuine trust. However, insincere and unstable persons as well as weakly collectively committed individualists may take part even in agreement-based cooperation. Thus such free riders, coat-turners, and dissidents may harm both the atmosphere and the joint project.

There is no foolproof recipe for finding trustworthy participants for cooperative ventures, nor for creating relationships of mutual respect where trust can grow and thrive. A good road to trust is to start by mutual evaluation of the features of the person and the situation and assurance of mutual respect. If this leads to predictive trust the parties take the risk and enter a cooperative action with each other, thus depending on the features of the others and the situation. If all goes well, and each party experiences that his rights are respected and that he respects the other's rights, he may believe that a relationship of mutual respect could be established. His predictive trust may grow into genuine trust when there are sufficiently many similar experiences. The crucial point is that only the experience of having one's rights respected, while respecting others' rights in a cooperative action can justify a belief of being a party of a respectful relationship. Prior to cooperation we can just estimate the likelihood of such a relationship getting established. Predictive trust can grow into genuine trust only through experience of mutual respect. Rational we-mode cooperation among sincerely committed persons and mutual belief about this satisfy the criteria for having a group member relationship of mutual respect, and rational persons should thus have genuine trust in each other vis-à-vis part-performances.

To sum up, our discussion has given us one central component of genuine trust, and the other assumptions needed for the existence of genuine trust are easily seen to be satisfied largely on the basis of rationality when the opportunities for action exist. Thus the conditions about ability, positive feeling, and accepting attitude are generally true of rational agents in the present kind of situation, although the participants of course cannot themselves intentionally make them true. The intention condition is true on the basis of the joint intention assumption of (*CWM*) in the we-mode case. In rational I-mode cooperation the cooperating parties have a reason to believe that the others have an intention to make their contribution, based on their expressions of intention or experience of their cooperative behavior or on some other ground.

5 Concluding summary

This paper is mainly about cooperation as a collective action in a group context, although the assumption of a group context need not be made in the case of I-mode cooperation. One of the central topics has been to clarify the different senses of cooperation that are relevant to group contexts, more specifically to contexts where the group members act as group members. A central distinction in the paper is between acting fully as a group member and acting as a private person within a group context. In the former case we speak of *we-mode* (and *we-perspective*) thinking and acting when the agent is committed as a group member to the main goals, standards, beliefs, etc. of the group. The cornerstone of the *we-perspective* is a general collectivity idea involved in colloquial slogans like "one for all and all for one", which can be technically explicated as Collectivity Condition applicable to collective attitudes, including goals, commitment, and actions. In cooperation the central intention expression is "We will achieve our goal". In the case of

acting as a private person for the group in a group context we speak of *I-mode pro-group* thinking and acting. Here the group member is only privately committed and has full control over his actions, which contrasts with the we-mode case where the authority over a member's action is with the group. In the I-mode case the central intention expression is "I will achieve my goal by cooperating with you".

The broadest notion of collective cooperative action is collective action performed on a "shared we-attitude" (see Tuomela and Bonnevier–Tuomela 1997 and R. Tuomela 2002a), excluding cases with conflict between either means actions. In the present paper some central cases of such broadly cooperative collective actions were considered. They satisfy the requirement than cooperating is necessarily intentional and involves the intention to cooperate.

A detailed analysis of I-mode cooperative action was given. In the two-person case the main idea is that the agents, say A_1 and A_2 , having the respective goals G_1 and G_2 are disposed to adjust their means actions for their own goals to the other's actions by which a participant, say A_1 , purports to achieve his goal. Thus, also the other agent, A_2 , will have his goal achievement promoted by A_1 's action. The means actions are either naturally dependent or are voluntarily taken by the participants to be dependent, and their goals become indirectly dependent because of this. When A_1 and A_2 are, and act as, members of a group with a certain ethos (constitutive goals, beliefs, norms, etc.) G_1 and G_2 are assumed to promote the ethos.

We-mode cooperative collective action can in simple cases be taken to consist of the participants' shared intended joint goal that they purport to achieve by their joint action. The joint action is a cooperative collective action in the we-mode when it promotes and is taken by the participants to promote the ethos of the group. This main idea is studied and clarified in detail in the paper. In our account it is essential that the participants, because of their jointly intended goal, have collectively committed themselves to the joint action and to each other to participate.

I-mode cooperation is cooperation based on the participants I-mode, "private" reasons while we-mode cooperation is based on we-mode reasons and thus on the adoption of the group perspective. Cooperation is a constitutive feature following from we-mode reasoning (in contrast to I-mode reasoning), and thus action based on we-mode reasoning will of necessity lead to some amount of cooperation. As we-mode cooperation is based on collective commitment and not only I-mode commitment, it is also more persistent than I-mode cooperation and helps to solve collective action dilemmas – even in one-shot cases of the Prisoner's Dilemma and analogous situations. The hard question is under what conditions the adoption of the we-mode is feasible. We do not properly consider this problem in the present paper, but emphasize the role of socialization in this matter.

Acting as a group member is argued to amount to cooperation towards the satisfaction of the ethos: a group member acts as a group member (in the core sense) if and only if he intentionally cooperates with respect to the group's ethos with the group members. This thesis divides into an I-mode and a we-mode version.

It is claimed in the paper that rational intentional cooperation conceptually, rationally, and psychologically requires at least predictive trust. On the other hand, it is argued that successful intentional cooperation can be taken to promote genuine trust based on mutual respect. We also concluded that we-mode cooperation among normally rational and sincerely collectively committed members, mutually believing this have a relationship of mutual respect and genuine trust in each other regarding part-performances.

Acknowledgments Maj Tuomela bears primary responsibility for Sect. 4, while the other sections are based on joint research. We wish to thank Robert Audi, Pekka Mäkelä, and Kaarlo Miller for comments.

References

- Bratman M (1992) Shared cooperative activity. *Philos Rev* 101:327–341
- Castelfranchi C, Falcone R (2001) Social trust: A cognitive model. In: Castelfranchi C, Yao-Hua T (eds) *Trust and deception in virtual societies*. Kluwer Academic Publishers, Dordrecht
- Conte R, Castelfranchi C (1995) *Cognitive and social action*. UCL Press, London
- Gilbert M (1989) *On social facts*. Routledge, London
- Gintis H (2003) Solving the puzzle of prosociality. *Ration Soc* 15:155–187
- Kollock P (1994) The emergence of exchange structures: An experimental study of uncertainty, commitment, and trust. *Am J Sociol* 100:313–345
- Mäkelä P, Tuomela R (2002) Group action and group responsibility. *Protosociology* 16:195–214
- McMahon C (2003) Shared agency and rational cooperation. *Nous* (forthcoming)
- Ostrom E, Walker J (eds) (2003) *Trust and reciprocity*. Russell Sage Foundation, New York
- Richerson P, Boyd R (2001) The evolution of subjective commitment to groups: A tribal instincts hypothesis. In: Nesse R (ed) *Evolution and the capacity for commitment*. Russell Sage Foundation, New York, pp 186–220
- Richerson P, Boyd R (2004) *Not by genes alone: How culture transformed human evolution*. The University of Chicago Press, Chicago
- Rosenschein J, Zlotkin G (1994) *Rules of encounter*. The MIT Press, Cambridge, MA
- Schutz A (1967) *The phenomenology of the social world* (1932), 1st edn. Northwestern University Press, Evanston, IL
- Sugden R (1993) Thinking as a team: Towards an explanation of nonselfish behavior. *Soc Philos Policy* 10:69–89
- Tuomela M (2003a) A collective's trust in a collective's action. *Protosociology* 18–19:87–126
- Tuomela M (2003b) The components of rational trust. In: Sintonen M, Ylikoski P, Miller K (eds) *Realism in action: Essays in the philosophy of the social sciences*. Kluwer Academic Publishers, Dordrecht, pp 168–180
- Tuomela M (2002) On rational trust. In: Meggle G (ed) *Social facts & collective intentionality*. Dr. Hänsel–Hohenhausen AG, Frankfurt am Main, pp 367–383
- Tuomela R (1984) *A theory of social action*, Synthese Library. Reidel Publishing Company, Dordrecht and Boston
- Tuomela R (1995) *The importance of us: A philosophical study of basic social notions*. Stanford University Press, Stanford
- Tuomela R (2000) *Cooperation: A philosophical study* (Philosophical Studies Series, vol 82). Kluwer Academic Publishers, Dordrecht
- Tuomela R (2002a) *The philosophy of social practices: A collective acceptance view*. Cambridge University Press, Cambridge
- Tuomela R (2002b) Joint intention and commitment. In: Meggle G (eds) *ocial facts & collective intentionality* (German Library of Sciences, Philosophical Research, vol 1), Dr. Hänsel–Hohenhausen AG, Frankfurt, pp 385–418
- Tuomela R (2003a) The we-mode and the I-mode. In: Schmitt F (ed) *Socializing meta-physics: The nature of social reality*. Lanham, Rowman and Littlefield, pp 93–128

- Tuomela R (2003b) We-intentions revisited. *Philos Stud* (forthcoming)
- Tuomela R (2004a) Cooperation and the we-perspective. Invited paper for the conference “Rationality and Commitment” held at the University of St. Gallen, May 2004 (forthcoming in its proceedings)
- Tuomela R (2004b) Joint action. Paper in the proceedings of the conference “Holistic epistemology and theory of action”, Leipzig, June 2004 (forthcoming)
- Tuomela R, Bonnevier–Tuomela M (1992) Social norms, tasks, and roles. Reports from the Department of Philosophy, University of Helsinki, No 1
- Tuomela R, Bonnevier–Tuomela M (1997) From social imitation to teamwork. In: Holmström–Hintikka G, Tuomela R (eds) *Contemporary action theory*, vol II: Social action. Kluwer Academic Publishers, Dordrecht Boston, pp 1–41
- Tuomela R, Miller K (1988) We-intentions. *Philos Stud* 53:115–137
- Tuomela R, Tuomela M (2003) Acting as a group member and collective commitment. *Protosociology* 18–19:7–65