

into and which are locked out of the theory/model. In other words you can't stipulate that the phenomena in a domain have just those features that your theory/model happens to be equipped to deal with. Moreover, you can't assume that what appears to be a good theory/model of a subdomain is automatically a good subtheory/partial model of the domain—essential features of the domain may have been excluded and the scope of the data may have been arbitrarily restricted. Improper simplification often occurs in semantics, for example, whenever a semantics for a "fragment" of English is given without any regard to whether it can coherently be extended to constructions not under consideration.

Any model of people's communicative behavior must cover people's ability to produce utterances with recognizable communicative intentions as well as their ability to recognize the communicative intentions behind other people's utterances. In my view there has been a tension in the ongoing effort to produce such a model: the guiding principles are too broad while the detailed implementations are too narrow. I agree with Thomason that general rules like Grice's maxims don't yield determinate predictions about concrete cases and, unless supplemented with much more detail, seem good for little more than ad hoc explanations. A model must be general enough to apply to a wide, indeed open-ended range of cases if it is to have any explanatory value, but, like Litman and Allen's, it must be detailed enough to make explicit the flow of information in the system being modeled. The dilemma is how to devise models that are neither too vague or otherwise uninformative nor too constrained or otherwise too specific.

References

- Bach, K. (1984). Default reasoning: Jumping to conclusions and knowing when to think twice. *Pacific Philosophical Quarterly* 65, 37–58.
- Bach, K. (1987a). *Thought and reference*. Oxford: Oxford University Press.
- Bach, K. (1987b). On communicative intentions: A reply to Recanati. *Mind and Language* 2, 141–154.
- Bach, K., and R. Harms (1979). *Linguistic communication and speech acts*. Cambridge, MA: MIT Press.
- Grice, H. P. (1957). Meaning. *The Philosophical Review* 66, 377–388.
- Grice, H. P. (1969). Uterer's meaning and intentions. *The Philosophical Review* 78, 147–177.
- Harnam, G. (1986). *Change in view*. Cambridge, MA: MIT Press.
- Perrault, C. R., and J. F. Allen (1980). A plan-based analysis of indirect speech acts. *American Journal of Computational Linguistics* 6, 167–182.
- Recanati, F. (1986). On defining communicative intentions. *Mind and Language* 1, 213–242.
- Schelling, T. (1960). *The strategy of conflict*. Cambridge, MA: Harvard University Press.
- Schiffer, S. (1972). *Meaning*. Oxford: Oxford University Press.
- Searle, J. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.
- Sperber, D., and D. Wilson (1986). *Relevance*. Cambridge, MA: Harvard University Press.
- Strawson, P. F. (1964). Intention and convention in speech acts. *The Philosophical Review* 73, 439–460.

Intentions in Communication
P. Cohen, J. Morgan & M. Pollack (eds)
MIT Press (1990)

Chapter 19

Collective Intentions and Actions

John R. Searle

This paper begins with an intuition, a notation, and a presupposition. The intuition is: Collective intentional behavior is a primitive phenomenon that cannot be analyzed as just the summation of individual intentional behavior, and collective intentions expressed in the form "we intend to do such-and-such" or "we are doing such-and-such" are also primitive phenomena and cannot be analyzed in terms of individual intentions expressed in the form "I intend to do such-and-such" or "I am doing such-and-such." The notation is: S(p). The "S" stands for the type of psychological state; the "p" stands for the propositional content; the content that determines the conditions of satisfaction. Like all such notations, it isn't neutral; it embodies a theory. The presupposition is: All intentionality, whether collective or individual, requires a preintentional Background of mental capacities that are not themselves representational. In this case that implies that the functioning of the phenomena represented by the notation requires a set of phenomena that cannot be represented by that notation.

The questions the paper addresses are: Is the intuition right? (It is denied by most of the authors I have read on the subject.) And if it is right, can it be made to fit the notation? How, if at all, can we capture the structure of collective intentions within that notation? And what role does the Background play in enabling us to function in social collectives? These questions are not innocent. The larger question they form a part of is: How far can the theory of intentional action in *Intentionality* (Searle 1983) be extended to become a general theory?

1 The Intuition

Let's start with the intuition. The first half of the intuition could hardly be wrong. It seems obvious that there really is collective intentional behavior as distinct from individual intentional behavior. You can see this by watching a football team execute a pass play or hear it by listening to an orchestra. Better still, you can experience it by actually engaging in some group activity in which your own actions are a part of the group action.

"we intend"

The problem is with the second half of the intuition, the idea that the collective behavior is somehow not analyzable in terms of individual behavior and that the collective intention is somehow not reducible to a conjunction of singular intentions. How, one wants to ask, could there be group behavior that wasn't just the behavior of the members of the group? After all, there isn't anyone left to behave once all the members of the group have been accounted for. And how could there be any group mental phenomenon except what is in the brains of the members of the group? How could there be a "we intend" that wasn't entirely constituted by a series of "I intend"s? There clearly aren't any bodily movements that are not movements of the members of the group. You can see that if you imagine an orchestra, a corps de ballet, or a football team. So if there is anything special about collective behavior, it must lie in some special feature of the mental component, in the form of the intentionality.

I want to build up to a characterization of the special form of collective intentionality by first trying to justify the first part of the original intuition.

Thesis 1

There really is such a thing as collective intentional behavior that is not the same as the summation of individual intentional behavior.

I said this seems obvious, but it is important to see how pervasive collective behavior is. It is by no means confined to human beings but rather seems to be a biologically primitive form of animal life. Studies of animal behavior are filled with accounts of cooperative behavior, but it does not take a specialist's knowledge to recognize. Consider two birds building a nest together, or puppies playing on a lawn, or groups of primates foraging for food, or even a man going for a walk with his dog. In humans, collective behavior typically involves language, but even for humans it does not invariably require language or even conventional ways of behaving. For example, I see a man pushing a car in the street in an effort to get it started, and I simply start pushing with him. No words are exchanged and there is no convention according to which I push his car. But it is a case of collective behavior. In such a case I am pushing only as part of *our* pushing.

Perhaps the simplest way to see that collective behavior is not the same as the summation of individual behavior is to see that the same type of bodily movements could on one occasion be a set of individual acts and could on another occasion constitute a collective action. Consider the following sort of example: Imagine that a group of people are sitting on the grass in various places in a park. Imagine that it suddenly starts to rain and they all get up and run to a common, centrally located shelter. Each person has the intention expressed by the sentence "I am running to the shelter."

But for each person, we may suppose that his or her intention is entirely independent of the intentions and behavior of others. In this case there is no collective behavior; there is just a sequence of individual acts that happen to converge on a common goal. Now imagine a case where a group of people in a park converge on a common point as a piece of collective behavior. Imagine that they are part of an outdoor ballet where the choreography calls for the entire corps de ballet to converge on a common point. We can even imagine that the external bodily movements are indistinguishable in the two cases, the people running for shelter make the same types of bodily movements as the ballet dancers. Externally observed, the two cases are indistinguishable, but they are clearly different internally. What exactly is the difference? Well, part of the difference is that the form of the intentionality in the first case is that each person has an intention that he could express without reference to the others, even in a case where each has mutual knowledge of the intentions of the others. But in the second case the individual "I intend"s are in a way we will need to explain, derivative from the "we intend"s. That is, in the first case, even if each person knows that the other people intend to run to the shelter and knows that the other people know that he intends to run to the shelter, we still do not have collective behavior. In this case at least, it seems no set of "intend"s, even supplemented with beliefs about other "I intend"s, is sufficient to get to the "we intend." Intuitively, in the collective case the individual intentionality, expressed by "I am doing act A," is derivative from the collective intentionality, "We are doing act A."

Another clue that collective intentions are different from a mere summation of individual intentions is that often the derived form of an individual intention will have a different content from the collective intention from which it is derived. We can see this in the following sort of example. Suppose we are on a football team and we are trying to execute a pass play. That is, the team intention, we suppose, is in part expressed by "We are executing a pass play." But now notice: no individual member of the team has this as the entire content of his intention, for no one can execute a pass play by himself. Each player must make a specific contribution to the overall goal. If I am an offensive lineman, my intention might be expressed by "I am blocking the defensive end." Each member of the team will share in the collective intention but will have an individual assignment that is derived from the collective but has a different content from the collective. Where the collective's is "We are doing A," the individual's will be "I am doing B," "I am doing C," and so on.

But supposing we got the characterization of the "I intend"s just right, couldn't we show how they add up to a "we intend"? I think not, and this leads to our second thesis:

Thesis 2

We-intentions cannot be analyzed into sets of I-intentions, even if intentions supplemented with beliefs, including mutual beliefs, about the intentions of other members of a group.

I think most philosophers would agree that collective behavior is a genuine phenomenon; the disagreement comes in how to analyze it. One tradition is willing to talk about group minds, the collective unconscious, and so on. I find this talk at best mysterious and at worst incoherent. Most empirically minded philosophers think that such phenomena must reduce to individual intentionality; specifically, they think that collective intentions can be reduced to sets of individual intentions together with sets of beliefs and especially mutual beliefs. I have never seen any such analysis that wasn't subject to obvious counterexamples, but let us try it out to see why it won't work. To have an actual sample analysis to work with, let us try that of Tuomela and Miller (1988), which is the best I have seen.

Leaving out various technical details, we can summarize their account as follows. An agent A who is a member of a group "we-intends" to do X if

1. A intends to do his part of X.
2. A believes that the preconditions of success obtain: especially, he believes that the other members of the group will (or at least probably will) do their parts of X.
3. A believes that there is a mutual belief among the members of the group to the effect that the preconditions of success mentioned in point 2 obtain.

This account is typical in that it attempts to reduce collective intentions to individual intentions plus beliefs. I, on the contrary, am proposing that no such reduction will work, that "we-intentions" are primitive. And I think it is easy to see what is wrong with the Tuomela-Miller account: a member of a group can satisfy these conditions and still not have a we-intention.

Consider the following situation. Suppose a group of businessmen are all educated at a business school where they learn Adam Smith's theory of the hidden hand. Each comes to believe that he can best help humanity by pursuing his own selfish interest, and they each form a separate intention to this effect; that is, each has an intention he would express as "I intend to do my part toward helping humanity by pursuing my own selfish interest and not cooperating with anybody." Let us also suppose that the members of the group have a mutual belief to the effect that each intends to help humanity by pursuing his own selfish interests and that these intentions will probably be carried out with success. That is, we may suppose that each is so well indoctrinated by the business school that each believes that his selfish efforts will be successful in helping humanity.

Now consider any given member A of the business school graduating class.

1. A intends to pursue his own selfish interests without reference to anybody else, and thus, he intends to do his part toward helping humanity.
2. A believes that the preconditions of success obtain. In particular, he believes that other members of his graduating class will also pursue their own selfish interests and thus help humanity.
3. Since A knows that his classmates were educated in the same selfish ideology that he was, he believes that there is a mutual belief among the members of his group that each will pursue his own selfish interests and that this will benefit humanity.

Thus, A satisfies the Tuomela-Miller conditions, but all the same, he has no collective intentionality. There is no we-intention. There is even an ideology, which he and the others accept, to the effect that there should not be a we-intention.

This case has to be distinguished from the case where the business school graduates all get together on graduation day and from a pact to the effect that they will all go out together and help humanity by way of each pursuing his own selfish interests. The latter case is a case of collective intentionality; the former case is not. Cooperative collective goals may be pursued by individualistic means, as is also shown by the following example. Suppose one of the members of a softball team loses his wallet at the game. Suppose the members reason that the chances of finding it are best if they each act separately; and each searches for the wallet in his own way, ignoring the others. They then set about in a coordinated and cooperative way to search for the wallet by acting with complete lack of coordination and cooperation. Unlike the original counterexample, these are genuine cases of collective behavior.

Could we avoid such counterexamples by construing the notion of "doing his part" in such a way as to block them? I think not. We are tempted to construe "doing his part" to mean doing his part toward achieving the *collective* goal. But if we adopt that move, then we have included the notion of a collective intention in the notion of "doing his part." We are thus faced with a dilemma: if we include the notion of collective intention in the notion of "doing his part," the analysis fails because of circularity; we would now be defining we-intentions in terms of we-intentions. If we don't so construe "doing his part," then the analysis fails because of inadequacy. Unless the we-intention is built into the notion of "doing his part," we will be able to produce counterexamples of the sort I have outlined above.

The reason that we-intentions cannot be reduced to I-intentions, even I-intentions supplemented with beliefs and beliefs about mutual beliefs, can be stated quite generally. The notion of a we-intention, of collective intentionality, implies the notion of *cooperation*. But the mere presence of I-intentions to achieve a goal that happens to be believed to be the same goal as that of other members of a group does not entail the presence of an intention to cooperate to achieve that goal. One can have a goal in the knowledge that others also have the same goal, and one can have beliefs and even mutual beliefs about the goal that is shared by the members of a group, without there being necessarily any cooperation among the members or any intention to cooperate among the members.

I have not demonstrated that no such analysis could ever succeed. I am not attempting to prove a universal negative. But the fact that the attempts that I have seen to provide a reductive analysis of collective intentionality fail for similar reasons—namely, they do not provide sufficient conditions for cooperation: one can satisfy the conditions in the analysis without having collective intentionality—does suggest that our intuition is right: we-intentions are a primitive phenomenon.

However, my claim that there is a form of collective intentionality that is not the product of some mysterious group mind and at the same time is not reducible to individual intentions has plenty of problems of its own, and we must now set about solving some of them. The most difficult problem we can put in the form: What exactly is the structure of we-intentions? We will not be in a position to answer that question until we answer a prior question about how we can reconcile the existence of collective intentionality with the fact that society consists entirely of individuals and no facts about any individual mental contents guarantee the existence of any other individuals. I believe it is facts such as these that have led people to believe that there must be a reduction of we-intentions to I-intentions.

Anything we say about collective intentionality must meet the following conditions of adequacy:

Constraint 1

It must be consistent with the fact that society consists of nothing but individuals. Since society consists entirely of individuals, there cannot be a group mind or group consciousness. All consciousness is in individual minds, in individual brains.

Constraint 2

It must be consistent with the fact that the structure of any individual's intentionality has to be independent of the fact of whether or not he is getting things right, whether or not he is radically mistaken about what is actually occurring. And this constraint applies as much to collective intentionality as it does to individual intentionality. One way to put this constraint is to say that the account must be consis-

tent with the fact that all intentionality, whether collective or individual, could be had by a brain in a vat or by a set of brains in vats.

These two constraints amount to the requirement that any account we give of collective intentionality, and therefore of collective behavior, must be consistent with our overall ontology and metaphysics of the world, an ontology and metaphysics based on the existence of individual human beings as the repositories of all intentionality, whether individual or collective.¹

Thesis 3

The thesis that we-intentions are a primitive form of intentionality, not reducible to I-intentions plus mutual beliefs, is consistent with these two constraints.

Actually, I think it is rather simple to satisfy these constraints. We simply have to recognize that there are intentions whose form is: We intend that we perform act A, and such an intention can exist in the mind of each individual agent who is acting as part of the collective. In cases like that of the football team each individual will have further intentional content, which in ordinary English he might express in the form "I am doing act B as part of our doing act A." For example, "I am blocking the defensive end as part of our executing a pass play." We need only note that all the intentionality needed for collective behavior can be possessed by individual agents even though the intentionality in question makes reference to the collective.

In the cases described above, if I am pushing only as part of our pushing, or if I am blocking the defensive end as part of our executing a pass play, the intentionality, both plural and singular, is in my head. Of course, I take it in such cases that my collective intentionality is in fact shared; I take it in such cases that I am not simply acting alone. But I could have all the intentionality I do have even if I am radically mistaken, even if the apparent presence and cooperation of other people is an illusion, even if I am suffering a total hallucination, even if I am a brain in a vat. Collective intentionality in my head can make a purported reference to other members of a collective independently of the question whether or not there actually are such members.

Since this claim is consistent with the brain in the vat fantasy, it is a fortiori consistent with each of our constraints. It is consistent with Constraint 2, because the brain in the vat formulation is just the most extreme form of stating this constraint; it is consistent with Constraint 1, because we are not required to suppose that there is any element in society other than individuals—that is, the supposition is entirely consistent with the fact that society is made up entirely of individuals. It is consistent with the fact that there is no such thing as a group mind or group consciousness,

because it only requires us to postulate that mental states can make reference to collectives where the reference to the collective lies outside the bracket that specifies the propositional content of the intentional state. The thought in the agent's mind is simply of the form "We are doing so and so."

Perhaps an uncomfortable feature of the analysis is that it allows for a form of mistake that is not simply a failure to achieve the conditions of satisfaction of an intentional state and is not a simply a breakdown in the Background. It allows for the fact that I may be mistaken in taking it that the "we" in the "we intend" actually refers to a we; that is, it allows for the fact that my presupposition that my intentionality is collective may be mistaken in ways that go beyond the fact that I have a mistaken belief. I do indeed have a mistaken belief if I have a collective intention that is not in fact shared, but on the proposed analysis, something further has gone wrong. Now, this does violate a very deep Cartesian assumption that we feel inclined to make. The assumption is that if I am mistaken, it can only be because one of my beliefs is false. But on my account, it turns out that I can not only be mistaken about how the world is but an even mistaken about what I am in fact doing. If I am having a hallucination in supposing that someone else is helping me push the car, that I am only pushing as part of our pushing, then I am mistaken not only in my belief that there is somebody else there pushing as well but also about what it is that I am doing. I thought I was pushing as part of our pushing, but that is not in fact what I was doing.

2 The Notation

I now turn to the notation. What exactly is the formal structure of collective intentionality? In order to state the structure of intentionality for singular actions. An action of, say, raising one's arm has two components: a "mental" component and a "physical" component. The mental component both represents and causes the physical component, and because the form of causation is intentional causation, the mental causes the physical of way of representing it. In ordinary English we can say: when I succeed, my trying to do something causes an effect of a certain type, because that is what I was trying to achieve. In the notation that I have found useful and perspicuous we can represent these facts, when the action is one of raising one's arm, as follows:

i.a. (this i.a. causes: my arm goes up) CAUSES: MY ARM GOES UP.

The expressions in lowercase letters represent the mental component. The type of intentional state is specified outside the bracket; in this case "i.a." stands for intention-in-action; and the expressions inside the bracket

represent the conditions of satisfaction, what must be the case if the state is to be satisfied. Where intentions are concerned, these conditions are causally self-referential; that is, it is part of the conditions of satisfaction that the state itself must cause an event of the type represented in the rest of the conditions of satisfaction. The expressions in capital letters on the right represent actual physical events in the world. If the i.a. is successful, then the action will consist of two components, a "mental" and a "physical" component, and the condition of satisfaction of the mental is that it should cause a physical event of a certain type. Since we are supposing it is successful, the above notation represents the fact that it does cause an event of that type. All of these facts are summarized in the above abbreviation.

I want the notation to seem absolutely clear, so I will write out a paraphrase in ordinary English, treating the whole expression as if it were a sentence instead of a diagram of the structure of an intention:

There is an intention-in-action that has as its conditions of satisfaction that that very intention-in-action causes it to be the case that my arm goes up; and all of that mental stuff really does cause it to be the case in the physical world that my arm goes up.

Now let us remind ourselves of how it works for a slightly more complex case involving a by-means-of relation. Suppose a man fires a gun by means of pulling the trigger. He has an intention-in-action whose content is that that very intention-in-action should cause the pulling of the trigger, which in turn should cause the firing of the gun. If the intention is satisfied, the whole complex event looks like this:

i.a. (this i.a. causes: trigger pulls, causes: gun fires) CAUSES: TRIGGER PULLS, CAUSES: GUN FIRES.

Once again, the expressions in lowercase letters represent the contents of the mind, and the expressions in capital letters represent what happens in the real world. Since we are assuming that the contents of the mind are satisfied in subsequent formulations, we can simply leave out the reference to the real world. If satisfied, the contents of the mind can be read off directly onto the world. Previously we introduced the colon, which is read (with appropriate adjustments) as "it to be the case that . . ." and enables us to convert the sentence or other expressions that follow into singular terms. Here we introduce the comma, which is read as "which" and converts the subsequent expressions into a relative clause. Thus, the stuff inside the parentheses in this example is to be read in English as follows:

This intention-in-action causes it to be the case that the trigger pulls, which causes it to be the case that the gun fires.

Now, let us apply these lessons to the study of collective behavior. To that end, let us look at another case.

Suppose Jones and Smith are engaged in a piece of cooperative behavior. Suppose they are preparing a hollandaise sauce. Jones is stirring while Smith slowly pours in the ingredients. They have to coordinate their efforts because if Jones stops stirring or Smith stops pouring, the sauce will be ruined. Each has a form of collective intentionality that he could express as "We are preparing hollandaise sauce." This is a collective intention-in-action and it has the following form:

ia. (this ia. causes: sauce is mixed).

Now the puzzle is, how does this collective intention cause anything? After all, there aren't any agents except individual human beings, and somehow intentional causation has to work through them and only through them. I believe one of the keys to understanding collective intentionality is to see that in general the by and by-means-of relations for achieving the collective goal have to end in individual actions. Thus, we might ask the cooks, "How are you preparing the dinner?" "Well," they might answer, "first by making the sauce; then by cooking the meat." But at some point somebody has to be in a position to say, for example, "I am stirring." In such cases the individual component of the collective actions plays the role of means to ends. Jones's stirring is the means to firing the sauce in the same sense that pulling the trigger is the means to firing the gun. Jones has an intentional content that we could express in English as:

We are making the sauce by means of me stirring.

And Smith has the intentional content:

We are making the sauce by means of me pouring.

From the point of view of each agent there are not two actions with two intentions that he is performing. Rather, just as in the gun case there is only one intention and one action—to fire the gun by means of pulling the trigger—so in the collective case each agent has only one intention that represents his contribution to the single collective action:

Jones: ia. (this ia. causes: ingredients are stirred).

Smith: ia. (this ia. causes: ingredients are poured).

But we still haven't solved our problem. In the case of the individual action there is a single intention that encompasses the by-means-of relations. I intend to fire the gun by means of pulling the trigger. One intention, one action. The relation of the means-intention to the overall intention is simply part-whole: the whole intention represents both the means

and the ends, and it does that by representing the by-means-of relation according to which one achieves the end by means of the means.

But how exactly does it work where the means is individual and the goal is collective? The answer to that question is not at all obvious. Let us try some possibilities. It is tempting to think that such intentions might contain collective intentionality right down to the ground, that there might simply be a special class of collective intentions and that is all that is needed. On this account, from Jones's point of view the intentionality is this:

collective ia. (this collective ia. causes: ingredients are stirred, causes: sauce is mixed).

But this "collectivist" or "socialist" solution can't be right because it leaves out the fact that Jones is making an individual contribution to a collective goal. If I am Jones, this account leaves it as mysterious how the collective intentionality can move my body. Surely one feels like saying, I personally have to intend to do something if the sauce is ever going to get mixed.

But the opposite view, according to which it is all individual intentionality, a "capitalist" or "individualist" solution, fares no better:

singular ia. (this singular ia. causes: stirred, causes: mixed).

This is unsatisfactory because it is consistent with there being no collective intentionality at all. I might stir in the knowledge that you were doing something that together with my stirring would produce the desired result without any collective intentionality. In short, this formulation is consistent with the claim that there is no such thing as collective intentionality; it is just an accumulation of individual intentionality, and that view we have already rejected.

Well, suppose we try to capture both the collective and individual components in the following way. Suppose we treat the collective intention as causing the singular intention:

collective ia. (this collective ia. causes: singular ia., causes: stirred, causes: mixed).

The feature of this analysis that makes me think it must be false is the fact that a separate ia. is in the scope of the collective ia. This would imply that the collective intention isn't satisfied unless it causes me to have a singular ia. And that can't be right, because my collective intention isn't an intention to make it be the case that I have a singular intention; it is the intention to achieve some collective goal for which my singular intention stands as means to end. A clue that this must be wrong is provided by the fact that it is quite unlike the case of ordinary singular action where my intention to fire the gun by means of pulling the trigger consists in only

one complex intention, not two intentions where one causes the other as part of its conditions of satisfaction. Of course, in the singular cases an intention can cause me to have a subsidiary intention, by practical reasoning. But even in such cases it doesn't necessarily have to cause the subsidiary intention in order to be satisfied. In the singular case there is just one intention in the agent's head: to fire the gun by means of pulling the trigger. Now why should there be two intentions in each agent's head in the collective case?

Well, let's try a new start. Let's ask intuitively what is going on. Intuitively, we are intentionally making the sauce and if I am Jones, my share is that I am intentionally stirring the ingredients. But what exactly is the relation between the collective and the individual intention? It seems to me it is exactly like the relation of the intention to pull the trigger and the intention to fire the gun: just as I fire the gun by means of my pulling the trigger, so we make the sauce by means of me stirring and you pouring. As far as my part is concerned, we intend to make the sauce by means of me stirring. But don't those have to be two separate intentions, one singular i.a. and one collective i.a.? No, no more than there have to be two separate intentions when I fire the gun by means of pulling the trigger. The real distinction between the singular and the collective case is in the type of intention involved, not in the way that the elements in the conditions of satisfaction relate to each other. The form of the intention in the singular case is to achieve goal B by way of doing means A. That is, it isn't just any old type of i.a., it is an achieve-B-by-means-of-A type of i.a. So we might think of the notation that represents this type of i.a. as containing two free variables, "A" and "B"; and these variables are then bound by clauses inside the brackets that function as nouns. What we are trying to say is that I have an achieve-B-by-means-of-A sort of intention whose content is that that-the-trigger-pulls-as-A causes it to be the case that-the-gun-fires-as-B. And we can represent this as follows:

i.a. B by means of A (this i.a. causes: A trigger pulls, causes: B gun fires).

Similarly, in the structure of collective action, there is only one (complex) i.a., and it isn't just any old type of i.a.; it is an achieve-collective-B-by-means-of-singular-A type of i.a. And when it comes to the notation, we bind those free variables in the representation of the type of intention by clauses functioning as singular noun phrases inside the brackets:

i.a. collective B by means of singular A (this i.a. causes: A stirred, causes: B mixed).

I am not sure this is the right analysis, but it does seem to be better than the three others we considered. It allows for both the collective and the

individual component in the agent's intentions. And it does so in a way that avoids making the paradoxical claim that the collective act causes the individual act. Rather, the individual act is part of the collective act. The intention to stir is part of the intention to mix by means of stirring in the same way that in the gun case the intention to pull is part of the intention to fire by means of pulling.

3 *The Presupposition*

But now the next question arises, what sort of beings are we that we have the capacity to form such intentions? Ultimately the answer to that has to be biological, but there is a more restricted sense of the question that we can still address: What general Background capacities and phenomena are presupposed by the sketch of collective intentionality? I have just given? The manifestation of any particular form of collective intentionality will require particular Background skills, the ability to stir or play football, for example. But are there any features of the Background that are general or pervasive (even if perhaps not universal) for collective behavior? I think there are, but they are not easy to characterize. They are the sorts of things that old-time philosophers were driving at when they said things like "Man is a social animal" or "Man is a political animal." In addition to the biological capacity to recognize other people as importantly like us, in a way that waterfalls, trees, and stones are not like us, it seems to me that the capacity to engage in collective behavior requires something like a pre-intentional sense of "the other" as an actual or potential agent like oneself in cooperative activities. The football team has the sense of "us against them" and it has that sense against the sense of the larger us of "teams playing the game"; the orchestra has the sense of "us playing in front of them" and it has that sense as part of the larger us of "participants in the concert." "But," one might object, "surely this sense of others as cooperative agents is constituted by the collective intentionality." I don't think so. The collective behavior certainly augments the sense of others as cooperative agents, but that sense can exist without any collective intentionality, and what is more interesting, collective intentionality seems to presuppose some level of sense of community before it can ever function.

It is worth noticing in passing that most forms of competitive and aggressive behavior are forms of higher-level cooperation. Two men engaged in a prizefight are engaged in a form of competition, but it is a form of aggressive competition that exists only within a higher-level form of cooperation. Each prizefighter has the intention to hurt the other, but they have these intentions only within the frame of the higher-order intention to cooperate with each other in engaging in a prizefight. This is the distinction between a prizefight and a case of one man simply assault-

ing another man in a dark alley. And what goes for the prizefight also goes for football games, business competitions, courtroom trials, and in many cases even armed warfare. For human beings, most social forms of aggressive behavior require higher-level cooperation. For one person even to insult another at a cocktail party requires an extremely sophisticated higher level of cooperation among the participants in the insult.

Not all social groups are engaged in goal-directed behavior all the time. Some of the time they are just, for instance, sitting around in living rooms, hanging out in bars, or riding on the train. Now the form of collectivity that exists in such cases isn't constituted by goal-directed intentionality, because there isn't any. Such groups are, so to speak, ready for action but they are not yet engaged in any actions (they have no collective intentions-in-action) nor are they planning any (they have no collective prior intentions). Nonetheless, they have the type of communal awareness that is the general precondition of collective intentionality.

On the basis of such preliminary reflections I want to advance the following thesis:

Thesis 4

Collective intentionality presupposes a Background sense of the other as a candidate for cooperative agency; that is, it presupposes a sense of others as more than mere conscious agents, indeed as actual or potential members of a cooperative activity.

Now, what is the argument for this thesis? I don't know of anything like a conclusive argument; nonetheless, the considerations that incline me to this view are something like the following. Ask yourself what you must take for granted in order that you can ever have or act on collective intentions. What you must suppose is that the others are agents like yourself, that they have a similar awareness of you as an agent like themselves, and that these awarenesses coalesce into a sense of us as possible or actual collective agents. And these conditions hold even for total strangers. When I go out of my door into the street to help push the stranger's car, part of the Background is that each of us regards the other as an agent and as a candidate to form part of a collective agent. But these are not in the normal case "beliefs," just as my stance toward the objects around me and the ground underneath me is that of their being solid, without my needing or having a special belief that they are solid; and just as my stance toward others is that of their being conscious agents, without my needing or having a special belief that they are conscious; so my stance toward others with whom I am engaged in collective behavior is that of their being conscious agents in a cooperative activity, without my needing or having a special belief to that effect.

I believe that if we could fully understand the Background sense of others as possible agents, we would see that certain attempts to understand the character of society must be wrong. It is tempting to think that collective behavior presupposes communication, that speech acts in conversation are the "foundation" of social behavior and hence of society. It is perhaps equally tempting to suppose that conversation presupposes collective behavior, that social behavior is the foundation of conversation and hence of any society in which communication plays an essential role. There is obviously something to be said for each of these views. But I am here suggesting that we cannot explain society in terms of either conversation in particular or collective behavior in general, since each of these presupposes a form of society before they can function at all. The biologically primitive sense of the other person as a candidate for shared intentionality is a necessary condition of all collective behavior and hence of all conversation. We can now conclude with:

Thesis 5

The notation, and hence the theory, of *Intentionality* together with a certain conception of the role of the Background can accommodate collective intentions and actions.

Note

1. Readers will recognize that these two constraints are close to "methodological individualism" and "methodological solipsism" as traditionally construed. I am anxious if possible to avoid sinking into the morass of the traditional disputes, so I am trying to present a version of these in which they can be construed as just commonsensical, pretheoretical requirements.

References

- Searle, John R. (1983). *Intentionality: An essay in the philosophy of mind*. New York: Cambridge University Press.
 Tuomela, Raimo, and Kaarlo Miller (1988). We-intentions. *Philosophical Studies* 53, 367-389.