

# What Is Joint Action? (JAM4, July 2011—DRAFT)

Stephen A. Butterfill  
<s.butterfill@warwick.ac.uk>

July 5, 2011

## Abstract

Several researchers have conjectured that joint action partially explains how sophisticated\* forms of cognition emerge in development or evolution (for example, ??). If any such conjecture is correct, what could joint action be?

The leading approach to characterising joint action focuses on shared intention (Bratman 1993, 2009). A shared intention is something that stands to a joint action roughly as an ordinary intention stands to an individual action. Given some widely held claims about rationality and knowledge, shared intention requires knowledge of intentions about intentions (Butterfill forthcoming). So shared intention already presupposes sophisticated\* (social) cognition. If the conjectures about evolution and development are true, not all significant cases of joint action can involve shared intention: if they did, abilities to engage in joint action would presuppose sophisticated\* forms of cognition and so could not explain how they are acquired in evolution or development.

An alternative approach to joint action starts from the claim, motivated by semantic considerations, that a joint action is an action with two or more agents (Ludwig 2007). An immediate objection to this claim is that, given any of several widely held views about action, events standardly offered as paradigm joint actions turn out not to be joint actions at all. This objection can be overcome by refining the claim that a joint action is an action with two or more agents. This alternative approach is necessary for characterising the joint actions that could explain how sophisticated\* forms of cognition emerge in development or evolution.

\* - yes, I do realise that the term ‘sophisticated’ isn’t very scientific; the talk will be more specific

## 1. The Challenge

I want to start with a challenge. The challenge is to explain the emergence, in evolution or development (or both), of sophisticated forms of social cognition.

By ‘sophisticated forms of social cognition’, I mean those which involve representing perceptions, knowledge states, intentions, beliefs and other propositional attitudes. In other words, the sort of social cognition measured by tasks such as the famous false belief task (?).

Why think that representing beliefs and other propositional attitudes is a sophisticated form of social cognition? Part of the reason is that representing propositional attitudes is hard in two senses. A body of evidence with humans suggests that representing beliefs and other propositional attitudes requires conceptual sophistication,

- for it has a protracted developmental course stretching over several years (??),
- its acquisition is tied to the development of executive function (??) and language (?), things which two-year-olds, scrub jays and chimpanzees are deficient in, and
- development of reasoning about beliefs in humans is be facilitated by explicit training (?) and environmental influences, such as siblings (??).

Representing propositional attitudes also appears to be cognitively demanding, requiring attention and working memory in fully competent adults (???).<sup>1</sup>

It makes sense that representing propositional attitudes should be conceptually and cognitively demanding. After all, these are states which form complex causal structures, have arbitrarily nest-able contents, and are individuated by their causal and normative roles in explaining thoughts and actions. If anything should take years to acquire and consume scarce cognitive resources it is surely cognition of states with that combination of properties.

So the challenge is to explain the emergence, in evolution or development (or both), of sophisticated forms of social cognition—that is, of cognition that involves representing perceptions, beliefs, intentions and other propositional attitudes.

---

<sup>1</sup> These findings, like the developmental findings mentioned in the previous paragraph, are open to challenge (e.g. ?). There is a wide variety of positions in this area; we provide indirect support for the view taken here elsewhere (??).

## 2. The Conjecture

So much for the challenge. How can we meet it?

Several researchers have conjectured that social interaction partially explains how sophisticated forms of cognition, including social cognition, emerge in development or evolution:

‘the unique aspects of human cognition ... were driven by, or even constituted by, social co-operation’ (?, p. 1).

‘perception, action, and cognition are grounded in social interaction’ (?, p. 103).

These conjectures concern social interaction or social co-operation generally. In fact these researchers focus’ is on joint action. And, like them, I want to focus on joint action.

So the conjecture I am interested is this:

The existence of abilities to engage in joint action partially explain how sophisticated forms of social cognition emerge in evolution or development (or both).<sup>2</sup>

Faced with this conjecture, you might be wondering is what joint action is. That is my question too:

If this conjecture is correct, what could joint action be?

Paradigm cases of joint action in development include:

- tidying up the toys together (Behne et al 2005)
- cooperatively pulling handles in sequence to make a dog-puppet sing (Brownell et al 2006)
- bouncing a ball on a large trampoline together (Tomasello and Carpenter 2007)
- pretending to row a boat together

---

<sup>2</sup> [\*don’t say because only matters given the later argument] Note that this is explanation *how*. Distinguish explanation in the sense of (we have sophisticated cognition in order that we can engage in joint action) from explanation in the sense of (abilities to engage in joint action explain how sophisticated forms of cognition emerge in evolution or development).

Philosophers' paradigm cases of joint action include painting the house together (Michael Bratman), lifting a heavy sofa together (David Velleman), preparing a hollandaise sauce together (John Searle), going to Chicago together (Christopher Kutz), and walking together (Margaret Gilbert).

So one constraint on any adequate answer to my question is that it must be consistent with the claim that these paradigm cases are indeed joint actions. It is these paradigm cases that given us an initial, intuitive anchor on what joint action is. Beyond this I take everything to be open to question.

### 3. Shared Intention

My question is, What could joint action be given that it grounds sophisticated social cognition? The usual way of thinking about joint action starts with the premise that all significant cases of joint action involve shared intention. For instance:

'I take a collective action to involve a collective intention.' (?, p. 5).

'The sine qua non of collaborative action is a joint [shared] goal and a joint commitment' (Tomasello 2008, p. 181)

'the key property of joint action lies in its internal component ... in the participants' having a "collective" or "shared" intention.' (?, pp. 444-5).

'Shared intentionality is the foundation upon which joint action is built.' (?, p. 381)

'it is precisely the meshing and sharing of psychological states ... that holds the key to understanding how humans have achieved their sophisticated and numerous forms of joint activity' (?, p. 369)

I note that Michael Bratman hasn't made any claim of this sort. He is much more cautious. Although one of his recent papers does claim that even very modest forms of joint action involve shared intention (?), strictly speaking, it is consistent with his view to hold that there are significant cases of joint action which do not involve shared intention. ('Significant' means significant relative to our conjecture, that is, having relevance to explaining how social cognition emerges.)

But what is shared intention? Intentions are not literally shared in the sense in which siblings can share a parent—on almost any account, to share an intention does not require there to be a single mental state with two or more subjects. And although intentions can literally be shared in the sense in which siblings can share genes, this sense of sharing is almost universally regarded as too weak.

In fact the term ‘shared intention’ is doubly metaphorical. For on almost any account, shared intentions are neither literally shared nor intentions.

This may be why there is **much disagreement on what shared intentions are**. Some hold that shared intentions differ from individual intentions with respect to the attitude involved (?; ?). Others have explored the notion that shared intentions differ with respect to their subjects, which are plural (?), or that they differ from individual intentions in the way they arise, namely through team reasoning (?), or that shared intentions involve distinctive obligations or commitments to others (?; ?). Opposing all such views, ?? argues that shared intentions can be realised by multiple ordinary individual intentions and other attitudes whose contents interlock in a distinctive way.

#### 4. Bratman on Shared Intention

For this talk I am simply going to adopt Bratman’s account. This account is generally taken as a point of departure by philosophers and some psychologists. It is theoretical coherent, no one has succeeded in identifying a valid objection to it in print, and I believe it captures one concept that is important for understanding joint action.

I don’t want to say too much about this account because you will shortly be hearing from the man himself. But I do need us to have the bare outlines of this account before us now.

Bratman’s account of shared intention has two parts, a specification of the functional role shared intentions play and a substantial account of what shared intentions could be. On the first part, Bratman stipulates that the functional role of shared intentions is to:

- (i) coordinate activities; (ii) coordinate planning; and (iii) provide a framework to structure bargaining (?, p. 99)

To illustrate: if we share an intention that we cook dinner, this shared intention will (iii) structure bargaining insofar as we may need to decide what to cook or how to cook it on the assumption that we are cooking it together; the shared intention will also require us to (ii) coordinate our planning by each bringing complementary ingredients and tools, and

to (i) coordinate our activities by preparing the ingredients in the right order.

Given this claim about what shared intentions are for, Bratman argues that the following three conditions are jointly sufficient<sup>3</sup> for you and I to have a shared intention that we J.

- ‘1. (a) I intend that we J and (b) you intend that we J
- ‘2. I intend that we J in accordance with and because of la, lb, and meshing subplans of la and lb; you intend that we J in accordance with and because of la, lb, and meshing subplans of la and lb
- ‘3. 1 and 2 are common knowledge between us’ (?, View 4)

On the substantial account given by Bratman, sharing intentions requires intentions about intentions (see Condition 2 in the quote above).<sup>4</sup>

Furthermore, each agent must know that the others have intentions about her own intentions; and this knowledge must be mutual (see Condition 3 above). So sharing an intention involves knowing that someone else knows that I have intentions concerning subplans of their intentions.

**There is not much theory of mind cognition that meeting these conditions doesn’t require.** So if we suppose that all joint action requires shared intention and that shared intention requires knowledge of others’ knowledge about our intentions concerning their intentions ... then there is no way that abilities to engage in joint action could play any role in either the evolution or the development of theory of mind cognition.

At this point we need to slow down. So far we have been discussing merely sufficient conditions for shared intention. For all we have said, there may be other states which realise the functional roles of shared intention and which do not require sophisticated social cognition.

However, I do not believe that this is really possible. According to Bratman’s functional characterisation, sharing intentions involves coordinating planning. There may be various psychological states whose functional role is to coordinate planning among agents. But, in paradigm cases of joint action as well as in the sorts of case that matter for understanding evolution and development, states which play this role will involve knowledge of others’ intentions because intentions are the basic elements of plans.

---

<sup>3</sup> In (1993), Bratman offers the following as sufficient and necessary conditions; the retreat to merely sufficient conditions occurs in Bratman (1999 [1997]) where he notes that “for all that I have said, shared intention might be multiply realizable.”

<sup>4</sup> Bratman emphasises this feature of the account: “each agent does not just intend that the group perform the [...] joint action. Rather, each agent intends as well that the group perform this joint action in accordance with subplans (of the intentions in favor of the joint action) that mesh” (Bratman 1992: 332).

Elsewhere Bratman says:

‘shared intentional agency consists, at bottom, in interconnected planning agency of the participants.’<sup>5</sup>

It seems likely that interconnected planning agency, almost however it is realised, will require knowledge of other’s intentions, and of their intentions concerning one’s own plans. If this is right, that shared intentional agency—and shared intention—requires full-blown theory of mind cognition at close to the limits of what humans are capable of.<sup>6</sup>

## 5. Transition

We have seen that, on the standard view,

1. All (significant) joint actions require shared intention.

and

2. Shared intention requires sophisticated theory of mind cognition.

Therefore:

3. Abilities to engage in joint action could play no significant role in explaining how sophisticated theory of mind cognition emerges.

This means that the standard view about joint action is incompatible with The Conjecture.

This is not to say that there is no explanatory role for shared intention at all. It may be that shared intention explains *why* we have sophisticated forms of theory of mind cognition. Perhaps we have these because they enable us to have shared intentions, which in turn enables us to cooperate effectively. I don’t know whether this is right, but maybe it is possible to explain *why* we have sophisticated forms of theory of mind cognition by appeal to joint action involving shared intention. But what we can’t do is explain *how* we acquire theory of mind cognition by appeal to anything involving shared intention. Shared intention presupposes, and so cannot explain, sophisticated theory of mind cognition.

One way around this would be to attempt to reject the second premise, as ? has attempted to do. I want to pursue the other alternative, which is to reject the first premise. In my view, the form of joint action we need to explain how theory of mind cognition emerges does not involve shared intention. But, as we will see, it is not trivial to show that there are significant forms of joint action that do not involve shared intention.

---

<sup>5</sup> ?, p. 11. See also ?, p. 5: ‘We begin with planning agents.’

<sup>6</sup> This claim is argued for in detail in ?

## 6. A joint action is an action with two or more agents

The question was what joint action could be given that it grounds social cognition. So far we only have a negative answer: if it grounds social cognition, not all joint action can involve shared intention. My aim in the rest of the talk, then, is to investigate whether it is possible to characterise joint action without shared intention.

I want to start with a claim from Kirk Ludwig's semantic analysis. A *joint action* is an action with two or more agents, as contrasted with an *individual action* which is an action with a single agent (? , p. 366).

To illustrate, suppose we are on a long, narrow footbridge over a river. We each individually sway from side to side with the intention of causing the bridge to wobble noticeably. While none of our efforts are individually sufficient, collectively they do cause the bridge to wobble noticeably. It seems plausible that making the bridge wobble is an action. Furthermore, it seems that we are each agents of this action. If that's right, making the bridge wobble is a joint action in Ludwig's sense. It is a joint action just because it is an action with more than one agent.

On this definition, it seems likely that joint action does not require shared intention in Bratman's sense.<sup>7</sup> It may be, for all that has been said so far, that none of us are aware of each other's efforts and falsely believe that we alone are responsible for the bridge's wobbling.<sup>8</sup> So there may be nothing which functions to coordinate our planning or to structure or bargaining. And the coordination of our actions, of our swaying, may be due partly to chance and partly to entrainment.

## 7. First Objection: too narrow

Superficially, then, Ludwig appears to have offered what we need: a characterisation of joint action on which not all joint actions involve shared intention. Unfortunately this definition has a fatal defect for our purposes (as Ludwig notes, \*page). For on standard views of action, an

---

<sup>7</sup> Ludwig (\*ref) makes this claim, but the example he uses to support it is, for reasons given below, not obviously a joint action on his definition.

<sup>8</sup> On this definition, joint action may also fail to involve shared intention not only as characterised by Bratman but also as characterised by almost any account. For, on almost any account, shared intention implies awareness of joint-ness, awareness of others' agency or awareness of others' states or commitments. The example suggests that none of these are required for an action to be joint in Ludwig's sense. Awareness of joint-ness is not required because we may doubt whether others are intentionally participating, and awareness of others' agency is not required because we may not be aware of the others at all; it is also clear that awareness of others' states or commitments may not be required.



action is a bodily movement or a trying. According to Donald Davidson,

‘our primitive actions, the ones we do not by doing something else, ... these are all the actions there are.’ (?, p. 59).

Davidson also argues that the only actions which are primitive are ‘mere movements of the body’ (?, p. 59). Others have suggested that primitive actions are not bodily movements but the tryings which precede them.

To illustrate, suppose that I unlock a door by turning a key, which I in turn achieve by moving my fingers. On Davidson’s view, which is the standard view, the action of my unlocking the door is the action of my moving my fingers. It may sound baffling but this is in some ways a very intuitive idea. My action, the difference that my agency makes in the world, is not constituted by the lock’s movement but only by the movements of my own body. Since my action causes the door to unlock, we can *describe* my action by saying that I unlock the door. But this doesn’t mean that my action is constituted in part by changes in the door; it means that my action is one that causes changes in the door. Whether the lock moves or not depends on things which I cannot directly influence; it is at most the movements of my body which are under my control when I act.

All of this means that, on Ludwig’s definition, what are taken to be paradigm cases of joint action would not be joint actions at all. Why is that? Consider the case where we cooperatively pull handles in order to make a dog-puppet sing. This is one of the paradigm cases of joint action from development. First I pull a lever and then you pull a lever. Here there are no bodily movements with more than one agent.<sup>9</sup> I move my body and then you move yours. So if Davidson is right that actions are bodily movements, then our making the dog-puppet sing does not

---

<sup>9</sup> Is this too quick? One might consider the composite of the our bodily movements to be an event, noting that this event involves two agents. If this event were an action, there would be an action of which we are both agents. This event is clearly distinct from both my pulling and your pulling. But could there really be a third action, one distinct from both my pulling and your pulling? Consider a parallel scenario, one that is as similar as possible to our making the dog-puppet sing except that a single agent, Coralie, pulls both levers. Unless we are prepared to allow that there are three actions in this scenario, we should not allow that there are three actions in the original scenario either. Now on some views of action it might be coherent to allow that there are indeed three (or more) actions here (\*refs). But, given Davidson’s claims, the composite of Coralie’s two pullings is not an action. This is because the event is not primitive: it consists of Coralie’s two pullings and nothing else and so, trivially, it can be analysed in terms of these actions. We should therefore draw the same conclusion about the composite of my pulling and your pulling: it is not an action. So on Davidson’s view, the events (if any) which involve both me and you as agents are not actions.

involve any action with more than one agent. It follows that on Ludwig's definition no joint action is involved.

And the same goes for all the paradigm cases of joint action in both development and in philosophy. In each case there are no bodily movements with more than one agent and therefore no joint actions.

Here's the challenge we face:

1. A joint action is an action with two or more agents.
2. Bodily movements 'are all the actions there are.' (Roth, p. 59).
3. What are taken to be paradigm cases of joint action do not involve bodily movements with more than one agent.

Therefore:

4. What are taken to be paradigm cases of joint action are not actually joint actions.

In short, Ludwig's definition is **too narrow**.

Note that this argument does not depend on the assumption that all bodily movements have at most one agent. This assumption is beside the point.<sup>10</sup> The point is that *my pulling*, this particular event, had only one agent (and your pulling likewise). Whether any actions have two or more agents is an issue the argument does not address.

Some philosophers broadly in agreement with Davidson hold that actions are tryings rather than bodily movements (e.g. Roth). If the above argument works given Davidson's position it also works given this position. In fact, the argument works given any position which respects two constraints: first, no action involves the movements of the dog-puppet; and, second, an event comprising two actions and nothing else is not an action. Whether my pullings is a bodily movement, a trying or anything else which stops short of including a dog-puppet's singing, the action is mine alone.

The above argument applies to cases that are widely taken to be paradigmatic joint actions. In making the sauce, you stir while I pour; in painting the house, you cover the outside while I do the inside; and in walking together you move your legs while I move mine. The above argument establishes that in these cases there are no actions with more than one agent. On the definition of joint action under consideration this means that there are no joint actions. This is why, given some widely accepted claims about action, the claim that a joint action is an action with two or more agents implies that many supposedly paradigm cases of joint action are not in fact joint actions.

---

<sup>10</sup> Any such assumption would require careful consideration, especially if Roth (Roth) is right that one agent can literally act on another's intention.

## 8. Avoiding the first objection

How should respond to the argument at the close of the previous section?

Well, why should we respond at all? Why not just bite the bullet and accept that supposedly paradigm cases are not actually joint actions? This might be worth considering if our project were to provide a semantic theory. But our project is to identify a definition of joint action that supports philosophical and scientific inquiry. And a definition on which few or no supposedly paradigm cases turn out actually to be joint actions is unlikely to serve that purpose. After all, the paradigm cases are among the things which anchor the phenomenon to be defined.

A different response would be to reject one of the argument's premises. The argument depends on the premise that bodily movements (or tryings) are all the actions there are (premise 2). One response would be to question this assumption (\*ref Chant). I want to take a different approach, one which does not require rejecting the standard views about action.

I think the objection is probably correct—probably Ludwig's definition is too narrow. I shall therefore revise Ludwig's definition of joint action (premise 1). I shall do this by appealing to the idea that there is an attenuated sense of agency in which individuals may be agents of events other than actions.

Pietroski proposes a sense in which an individual can be the agent of an event which is not an action. His proposal has two parts. First, there is a relation among events, *grounding*. He stipulates that

‘event  $D$  grounds  $E$ , if:  $D$  and  $E$  occur;  $D$  is a (perhaps improper) part of  $E$ ; and  $D$  causes every event that is a proper part of  $E$  but is not a part of  $D$ .’ (? , p. 81)

Pietroski's intention is that the toppling of a line of ten dominoes should be grounded by the toppling of the first domino. The definition of grounding may need modification if it is to fulfil this intention. For suppose that the toppling of the first two dominoes is an event, call it  $F$ . Then, since  $F$  is a proper part of the toppling of the whole line and not a part of the toppling of the first domino, the above definition entails that the toppling of the first domino can only ground the toppling of the whole line if the toppling of the first domino causes  $F$ . But since the toppling of the first domino is a part of  $F$ , on many standard accounts of causation the former will not cause the latter. (More precisely, consider two principles. First, no event causes itself. Second, where one event causes another, the first event also causes any events which are part of the second. These principles jointly imply that the toppling of the first domino does not cause  $F$ .) If this is right, the above definition entails that the toppling of

the first domino does not ground the toppling of whole the domino line, contrary to what was intended.

We can overcome this potential objection (and related complications) by modifying the definition of grounding. Let us say that two events *overlap* just if a (perhaps improper) part of one is a (perhaps improper) part of the other. More generally (this will be useful later), two or more events *overlap* just if any (perhaps improper) part of one of these events is a (perhaps improper) part of any of the other events. Then:

**singular grounding revised** Event  $D$  *grounds*  $E$ , if:  $D$  and  $E$  occur;  $D$  is a (perhaps improper) part of  $E$ ; and  $D$  causes every event that is a part of  $E$  but does not overlap  $D$ .

For instance, suppose Andy paints the house. As we have seen, on Davidson's and other standard views, Andy's actions are not events which involve paint adhering to the walls. But it is plausible that his actions (whatever exactly they are) ground a larger painting episode, which starts with his actions and end with paint sticking to the house.

The second part of Pietroski's proposal is along these lines (we ignore some complications not relevant for present purposes): for any event, whether or not it is an action, to be an agent of that event is to be an agent of an action which grounds it (p. 82). This provides an attenuated sense in which an individual can be an agent of an event even if the event is not an action. The proposal also incorporates prior truths about agency since every event grounds itself. Given standard views about what actions are, there is a strong argument for accepting Pietroski's proposal. The argument is simply that, in ordinary thinking about action, people do identify individuals as agents of events which are not actions, and their doing so appears to serve practical purposes.

One way to avoid the first objection to Ludwig's definition of joint action is to revise it by adopting a notion of agency which is attenuated along the lines indicated by Pietroski. To make this work we first have to generalise his definition of grounding so that an event can be grounded by any number of events, not just one.

**plural grounding** Events  $D_1, \dots, D_n$  *ground*  $E$ , if:  $D_1, \dots, D_n$  and  $E$  occur;  $D_1, \dots, D_n$  are each part of  $E$ ; and every event that is a part of  $E$  but does not overlap  $D_1, \dots, D_n$  is caused by some or all of  $D_1, \dots, D_n$ .

For example, to return to the earlier illustration, my and your pullings on the handles ground our making the dog-puppet sing, the event which starts with the pullings and ends with the singing.

The generalised definition of grounding has the consequence that if events  $D_1$  and  $D_2$  ground  $E$  and  $D_1$  causes  $D_2$ , then, given that causation

is a transitive relation,  $D_1$  alone will also ground  $E$ . This and other ways in which the grounding relation may not be unique call for caution in generalising Pietroski's statement about agency. In particular, we need to allow for the possibility that more than one set of actions may ground an event. This can be done as follows:

For an individual to be among the agents of an event is for there to be actions  $a_1, \dots a_n$  which ground this event where the individual is an agent of one or more of these actions.

So where some actions ground an event, all the agents of those actions are agents of the event; and only agents of actions which ground the event are agents of the event.

How does this proposal apply in the case where you and I pull handles in sequence to make a dog-puppet sing? Consider the whole episode, the episode encompassing our two pullings and the dog-puppet's singing. Each of our pullings is an action of which we are agents and these two pullings ground the whole episode. So on this proposal you and I are agents of the event of the episode of the dog-puppet's singing. So, on this proposal, our making the dog-puppet sing is a joint action.

The proposal is also consistent with the view that supposedly paradigm cases of joint action, such as two people's painting a house together, really are joint actions. This is because in all paradigm cases, there are events with two or more agents. The proposal thus allows us to combine two claims about paradigm cases of joint action: first, that they do not involve *actions* with more than one agent (as standard views about action require); and, second, that they do involve *events* with more than one agent.

Accordingly one way of avoiding the first objection to Ludwig's definition of joint action is to revise it in line with this proposal. On the revised definition, a joint action is an event with more than one agent.

## 9. Second objection: too broad

The revised definition of joint action clearly avoids the first objection, for it no longer fails to classify paradigm cases as joint actions. But in revising the definition to avoid this objection we have left it open to a converse objection. As we shall see, whereas the original simple definition was too narrow the revised definition appears to be too broad, classifying as joint actions events which arguably should not be so classified.

Consider an example. Nora and Olive killed Fred. Each fired a shot. Neither shot was individually fatal but together they were deadly. An

ambulance arrived on the scene almost at once but Fred didn't make it to the hospital. On the revised simple definition, this event is a joint action just because Nora and Olive are both agents of it. Now suppose that Nora and Olive have no knowledge of each other, nor of each other's actions, and that their efforts are entirely uncoordinated. We might even suppose that Nora and Olive are so antagonistic to each other that they would, if either knew the other's location, turn their guns on each other. The event of their killing Fred is nevertheless a joint action on the revised simple definition. But unless one thinks of the central event of *Reservoir Dogs* (\*ref) as a joint action, this is likely to seem counterintuitive.

This objection is not decisive in the way that the first objection was. We should be cautious about moving from the premise that certain events are not intuitively joint actions to the conclusion that they are not actually joint actions. Equally, it may be a mistake to expect a *definition* of joint action to pick out only cases which are theoretically significant to certain explanatory projects.

So I do think we can say is that the revised definition is incorrect. What we can say is just that it doesn't sufficiently identify the form or forms of joint action relevant to understanding the Conjecture that joint action explains how sophisticated social cognition emerges.

## 10. Goal-directed joint action

What is missing from the revised definition of joint action (according to which a joint action is an event with two or more agents)? All of the cases of joint action that serve as paradigms in development or philosophy are *goal-directed joint actions*. That is, they are cases where the event taken as a whole is directed to a goal.

To illustrate, return to Nora and Olive's killing of Fred. Nora and Olive might not have been acting individually. Nora, asked about their action, might insist, 'the goal of our action was not to kill Fred but Leslie'. (Note that this statement concerns the goal of an action and does not explicitly mention the agents' intention.) What is it for the goal of Nora and Olive's action to be that of killing Leslie? Among all the actual and possible outcomes of their action, what distinguishes Leslie's killing as a goal to which their joint action is directed?

More generally, **what is the relation between a joint action and the goal (or goals) to which it is directed?** In answering this question it is tempting to appeal to shared intention. A shared intention functions to coordinate agents' activities and involves states which represent an outcomes. Where joint action involves shared intention, what makes it the case that an outcome is one to which the joint action is directed is the

fact that the shared intention which is coordinating the agents' actions involves states which represent this outcome.

As explained earlier, we cannot appeal to shared intention in characterising forms of joint action which might explain the emergence of social cognition. So our challenge is to find a way of understanding how joint actions can be goal-directed which does not involve shared intention.

### 10.1. Detour: Goals are not intentions

Before I can address this challenge I need to make a detour. It is quite common not to distinguish goals from intentions, or goal-directed action from intentional action. Failure to make this distinction would be fatal for my purposes. It matters for me that these are distinct and, also, that representing goals is likely to be less cognitively and conceptually demanding than representing intentions. So let me pause here to explain, by way of background, the distinction between goals and intentions.

The term 'goal' has been used for outcomes, possible or actual, to which actions might be directed. This use occurs in phrases like 'the goal of our struggles'. The same term has also been used, perhaps improperly, for psychological states; it is in this second sense that agents' goals might cause their actions. I use the term 'goal' in the former sense only. In my terms, a goal is an outcome to which an action might be directed. So a goal is not a psychological state; indeed, it is not a state of an agent at all.

A basic question about action—about ordinary individual action—is

What is the relation between an action and the goal (or goals)  
to which it is directed?

To illustrate, any action typically has many outcomes. Only very few of these will be goals to which the action was directed. For example, I grab Isabel's hands and swing her around. One outcome of my action is that she laughs. Another outcome of the same action that I break some glass. It's possible that only one of these outcomes was among the goals to which my action was directed. So the question could be put by asking what distinguishes the outcomes which are goals from all the others.

A standard answer is that intentions (or other goal-states) relate actions to the goals. An intention both causes my action and represents a possible or actual outcome of the action. In this way the intention relates my action to the goal.

Importantly this is not the only way that actions might be related to the goals to which they are directed. An alternative is possibility is that goals are related to actions by being their teleological functions. There's some controversy over exactly how to characterise teleological

functions, but here's an approach that will do for now. **For an outcome to be the teleological function of an action means that (i) in the past, actions of this type have caused outcomes of this type; (ii) this action happens now in part because actions of this type caused outcomes of this type in the past** To illustrate, suppose that I have swung Isabel in the past and this has caused her to laugh, and that I swing Isabel now in part because doing so has caused her to laugh in the past. Then on the simple account, making Isabel laugh is a teleological function of my action. So it's arguably possible for actions to be related to their goals by virtue of teleological functions, not just intentions.

It doesn't matter for what follows whether you agree with me that some actions really are related to their goals by virtue of teleological functions. All that matters is that coherent for someone to think that this is a possibility. This means that **Someone might be able to assign goals to an agent's actions without being able to assign intentions.**

Let me put this another way. It is coherent to suppose that someone's mindreading abilities might extend to being able to represent the goals of actions but not the intentions of agents.

I mention this because it is useful for understanding how there might be joint action without shared intention. To put the idea metaphorically, I am going to suggest that **joint action sometimes involves sharing goals rather than intentions.**

## 10.2. A series of notions ...

So, as I said before, our challenge is to find a way of understanding how joint actions can be goal-directed which does not involve shared intention. I'll meet this challenge this by giving you a series of increasingly elaborate notions. Each notion describes a relation between multiple agent's actions and an outcome. The first notion is that of distributive goal.

## 10.3. Distributive Goals

An outcome is a distributive goal of multiple agents' activities just if this outcome is a goal to which each agent's activities are individually directed and it is possible for all agents (not just any agent, all of them together) to succeed relative to this goal.

To illustrate, one dark night two communists each independently intend to paint a large bridge red. Because the bridge is large and they start from different ends, they have no idea of the other's involvement in their project until they meet in the middle. Although their intentions were simply to paint the bridge and did not explicitly involve agency at



all, they both succeed in painting the bridge. As this illustration suggests, it is possible to have a distributive goal without having any knowledge of, or intentions about, other agents or other actions.

Where multiple agents' activities have a distributive goal there is a sense in which their activities are directed to a goal. But this may amount only to each agent's activities being individually directed to that goal. For significant cases of joint action we need a richer notion, one that relations joint actions to goals without this being only a matter of each agent's activities being individually directed to the goal.

#### 10.4. Collective Goals

For an outcome to be a *collective goal* of a joint action, or of multiple agents' activities, three conditions must be met:

1. the outcome is a distributive goal of the agents' activities
2. the agent's activities are coordinated; and
3. coordination of this type would normally facilitate occurrences of outcomes of this type.

These features constitute what I call a *collective goal*. Any outcome with these three features is a collective goal of the joint action.

The communist bridge painters that I mentioned earlier, their activities do not have a collective goal because they are not coordinated. Examples of activities that typically have collective goals include uprooting a small tree together and tickling a baby together to make it laugh.

The notion of a collective goal assumes that of coordination. This should be understood in a very broad sense. When two agents between them lift a heavy block by means of each agent pulling on either end of a rope connected to the block via a system of pulleys, their pullings count as coordinated just because the rope relates the force each exerts on the block to the force exerted by the other. In this second case, the agents' activities are coordinated by a mechanism in their environment, the rope, and not necessarily by any psychological mechanism. By invoking a broad notion of coordination and invoking coordination of activities rather than of agents, the definition of collective goal avoids direct appeal to psychological states.

The word 'collective' in 'collective goal' should not be understood to imply that the agents involved constitute a collective in any social sense. Nor does having a collective goal imply that the agents think of themselves as having a collective goal. The use of 'collective' and 'distributive' reflects (but does not exactly match) the use of these terms

in literature on plural quantification.<sup>11</sup> The point is that for multiple agents' goal-directed activities to have a certain collective goal is not equivalent to each of their activities separately having that goal.

Where a joint action has a collective goal there is a sense in which, taken together, the activities are directed to the collective goal. It is not just that each agent individually pursues the collective goal; in addition, there is coordination among their activities which plays a role in bringing about the collective goal. We can put this in terms of the direction metaphor. Any structure or mechanism providing this coordination is directing the agents' activities to the collective goal. The notion of a collective goal provides one way of making sense of the idea that joint actions are goal-directed actions.

### 10.5. Shared Goals

Some joint actions involve potentially novel goals and are voluntary with respect to their jointness. For these cases, coordination of the agents' activities must involve psychological components. What is the minimum we must add in order to characterise this sort of joint action? I don't think we need shared intention. What we need to suppose is just that the agents are aware of their activities as having a distributive goal and expect that their actions will succeed only in concert with others' efforts.

This is captured by a third and final notion, the shared goal. For an outcome to be a *shared goal* of two or more agents' activities is for these all to be true:

1. the outcome is collective goal of their activities;
2. and the coordination is explained in part by the fact that:
  - (a) each agent expects each of the other agents to perform activities directed to the goal; and
  - (b) each agent expects the goal to occur as a common effect of all their goal-directed actions.

In favourable circumstances this simple pattern of goals and expectations would be sufficient to coordinate the agents' activities in bringing about this outcome.

---

<sup>11</sup> See Linnebo (?) for an introduction to plural quantification. On some views, the predicate in 'The goal of their activities was to lift this block' could be interpreted as either distributive and collective. On the distributive reading, the truth of the sentence is entailed by the truth of 'For each of their activities, the goal of that activity was to lift this block'. On the collective reading this entailment might not hold.

To illustrate, my goal is to lift this table, and I anticipate that your actions will also be directed to this goal and that the table's moving will occur as a common effect of our efforts; and your goals and expectations mirror mine. Our activities could be coordinated around the table's movement in virtue of this interlocking pattern of goals and expectations.

Although I have labelled this pattern of goals and expectations a shared goal, I'm nervous about invoking the term 'sharing' because this has lots of romantic associations. And of course shared goals are not literally shared. You can't share a goal—or an intention, for that matter—in the sense that you can share a parent with a sibling. So talk about sharing is just a colourful metaphor; what it amounts to in this case is just that each agent has expectations about others' goals and the efficacy of their actions.

## 11. Collective Goals vs. Shared Intentions

My aim in introducing the notions of collective and shared goals was to understand how joint actions could be related to their goals in the absence of shared intention. My claim is that the notion of a collective goal can be used to identify one way in which joint actions are related to their goals. Forms of joint action which might ground social cognition are the forms characterised by collective and shared goals. It is these notions rather, than that of shared intention, that we need to understand how social cognition emerges.

This is not an objection to the claim that shared intentions exist, and that fully understanding joint action involves understanding shared intention.

It is clear, I think, that agents can have collective and shared goals without having shared intentions. I now want to argue that the converse is also possible. That is, agents sometimes share intentions without their activities having any corresponding collective goal. To illustrate, suppose that four merchants get together and agree to fix their prices with the shared intention that they will each become rich enough to retire by the end of the year. The year turns out to be good for trade and the merchants' shared intention is realised. As it happens these merchants are excellent traders but poor strategists. Their profits would have been even larger if they had not colluded in fixing prices. In fact their price-fixing strategy was so flawed that coordination of this type could not normally have a positive effect on profits. So their attempts at coordination hindered rather than facilitated the realisation of their shared intention. Because the existence of a collective goal requires coordination of agents' activities to be of a type instances of which would normally facilitate

occurrences of goals of this type, no collective goal of the merchants' activities as here described corresponds to their shared intention. (Of course their activities are directed to the goal of collective enrichment *in some sense*, just not in the sense identified by the above characterisation of the notion of a collective goal.)

The possibility that shared intentions exist without there being corresponding collective goals is not a superficial feature of the way we have defined collective goal. Where there are shared intentions, one or more propositional attitudes serve to link the agents and their goals with any means of coordination. For example, the merchants described above believe, falsely, that coordinating their pricing will facilitate fulfilment of their shared intention. Apart from their beliefs and other propositional attitudes, nothing appropriately connects their coordination to their activities and their goal. In characterising collective goals our aim is to better understand how joint actions are related to their goals when such propositional attitudes are absent. In the case of collective goals, then, the link between a goal and a means of coordination can only involve facts, not beliefs or expectations, about its efficacy. This is why, on standard accounts of shared intention, it is possible for agents to act on shared intentions without there being any corresponding collective goal.

In short, some joint actions involve collective and shared goals with no shared intentions, and that other joint actions involve shared intentions with no corresponding collective or shared goal. This suggests that the revised simple definition of joint action, according to which a joint action is an event with two or more agents, cannot be replaced with a definition involving collective goals. If we attempted to do this, we would exclude some activities involving shared intention. For this reason it seems to me plausible that, as far as *defining* joint action is concerned, we should stick with that definition. Although it is too broad in encompassing many cases which are neither intuitively joint actions nor of any obvious theoretical significance, it is a useful starting point for at least two distinct ways of narrowing focus, one based on collective and shared goals, the other on shared intention.

## 12. Conclusion So Far (I may stop here)

These three notions—shared goal, collective goal and distributive goal—identify three ways in which a joint action could be related to its goal. They provide a foundation for characterising forms of joint action without shared intention.

My question was which joint actions ground social cognition. The negative answer was, Not those which involve shared intention. Where

joint action involves shared intention it generally requires sophisticated social cognition, social cognition which is close to the limits of what human adults are capable of. Forms of joint action which involve shared intention *presuppose* sophisticated social cognition and so cannot *explain how* it emerges in either evolution or development.

But, positively, there are also forms of joint action which do not involve shared intention. Some of these require no social cognition at all; others require only minimal theory of mind cognition, such as the ability to identify the goals of others' actions. These are the forms of joint action which are needed to make sense of the Conjecture that joint action explains how sophisticated forms of social cognition emerge.

Forms of joint action which are goal-directed but involve nothing like shared intention have been neglected by philosophers, perhaps partly because they are, or are thought to be, too simple to present conceptual puzzles. This is a mistake. To understand the cognitive bases of abilities to engage in joint action or their evolution or development, the fact that a conception of joint action presents conceptual puzzles is no virtue. It may be better to start with the simplest possible notions, such as those of distributive and collective goal, and add the minimum required to further demarcate the category of interest in any given inquiry. Of course, shared intention is important for fully understanding joint action. But to focus only on shared intention is to focus on the most sophisticated case—it is to focus on something that comes only at the end of evolution and development as we currently know them. To understand which joint actions ground social cognition, it is collective and shared goals that we need to focus on.

### 13. Further Question (I may decide to cut this)

I started with a Challenge, which was to explain how sophisticated forms of social cognition emerge in evolution or development, and a Conjecture, which was that joint action plays a role in explaining this. Do the notions of collective and shared goals help us to understand how joint action could explain the emergence of sophisticated forms of social cognition?

How might abilities to engage in the sorts of joint action which involve collective or shared goals be involved in the evolution or development (or both) of sophisticated social cognition?

My suggestion concerns identifying the goals of actions. It is one thing to have a general ability to recognise goals and quite another to be able to recognise the goals of this particular activity. To illustrate, consider Hare and Tomasello (2004). The pictures stand for what par-

ticipants in this experiment saw. The participants were chimpanzees. The question was whether the participants would be able to work out which of two containers contained a reward. On the left there is a chimpanzee who is trying but failing to reach for the reward. Chimpanzees have no problem getting the reward in this case, suggesting that they understand the goal of the failed reach. On the right there is a human pointing to the reward location. Chimpanzees do not reliably get the reward in this case, suggesting that they fail to understand the goal of the pointing action. This is one illustration of how identifying the goals of particular actions can be difficult.

Some of the most plausibly unique aspects of human cognition depend on our abilities to recognise the goals of novel behaviours involving tools and gestures. **In particular, communicative actions depend on our abilities to recognise as the goals of behaviours goals which the behaviours can serve only because we recognise the behaviours as serving those goals [the Gricean circle].** It is here that I think joint action can play a role in explaining aspects of human cognition. My suggestion will be, roughly, that abilities to engage in joint action provide a route to knowledge of others' goals which is distinct from ordinary third-person interpretation.

To explain this suggestion in detail I first need to describe some reasons why it can be hard to identify the goals of particular actions ...

### 13.1. The Problem of Opaque Means

Suppose you cannot gain knowledge of the current goals of another's actions through linguistic communication.

Suppose the goal is relatively novel (there are no stereotypical indications).

In this situation we cannot generally do better to work backwards from her behaviours to her goals: we determine which outcomes her behaviour is likely to bring about and then suppose that her goal is to bring about one or more of these outcomes (Dennett 1991).

But this method doesn't work when:

- \* we don't know which outcomes the observed behaviour is likely to bring about;
- \* we, or the agent under observation, have false beliefs relevant to which outcomes this behaviour is likely to bring about; or

So one obstacle to identifying the goals of particular actions is the problem of opaque means ...

Another problem involves false belief but I won't mention that here.<sup>12</sup>

### 13.2. The your-goal-is-my-goal route to knowledge

How could abilities to engage in joint action provide us with knowledge of others' goals? The intuitive idea I started with was this: if you're engaged in joint action with me, it's easy for me to know what your goal is ... because your goal is my goal.

This intuitive idea isn't quite right as it stands. For to be engaged in joint action requires that I already have expectations about the goals of your behaviours. So engaging in joint action presupposes rather than explains knowledge of others' goals. Or so it seems.

But there is a way around this. For there are various cues that you can give me which signal that you are about to engage in joint action with me. Seeing me struggling to get my twin pram on to the bus, you grab the front wheels and make eye contact, raising your eyebrows and smiling. In this way you signal that you both disposed to help and are about to engage in joint action with me. This makes it trivial for me to know what the goal of your behaviour is: your goal is my goal, to get the pram onto the bus.

My suggestion, then, is that the following inference characterises a route to knowledge of others' goals:

1. We are about to engage in some joint action<sup>13</sup> or other (for example, because you have made eye contact with me while I was in the middle of attempting to do something).
2. I am not about to change my goal.

Therefore:

3. The others will each individually perform actions directed to my goal.

---

<sup>12</sup> The interdependent roles of beliefs, desires and goals in producing action mean there will be many cases where observed behaviours are compatible with different ascriptions. To illustrate, consider Maya who is tidying shapes into two boxes. She mostly puts the squares into Leo's box. This indicates that the goal of her activity may be to put the squares into Leo's box, but it also leaves open the possibility that her goal is to put the squares into Charlie's box and she has a false belief about the owners of the boxes. In general, non-communicative behaviour indicates what an agent's goals are only given assumptions about her beliefs, and it indicates what her beliefs are only given assumptions about her goals (Davidson 1974 [1984]).<sup>18</sup> In some everyday situations this interdependence is a practical problem for knowing what others are doing. We could solve the problem if we had some way of getting at an agent's goals independently of knowing what she believes.

<sup>13</sup> \*What notion of joint action is needed here? Any will do as long as it involves distributive goals.

Call this the ‘your-goal-is-my-goal’ route to knowledge. To say that this inference characterises a route to knowledge implies two things. First, in some cases it is possible to know the three premises, 1–2, without already knowing the conclusion, 3. Second, in some cases knowing the two premises puts one in a position to know the conclusion. I take both points to be true.

The your-goal-is-my-goal route to knowledge is characterised by an inference. However, exploiting this route to knowledge may not require actually making the inference or knowing the premises. Depending on what knowing requires, it may be sufficient to believe the conclusion because one has reliably detected a situation in which the premises of the inference are true without necessarily being able to think of this as a situation where the premises are true.

### 13.3. Application

I want to suggest that your-goal-is-my-goal might give us a way to understand how joint action facilitates a transition from a simple understanding of goals to an understanding of communicative intent.

I already mentioned Hare and Call’s (?) experiment which contrasts pointing with a failed reach as two ways of indicating which of two closed containers a reward is in. Chimps can easily interpret a failed reach but are stumped by the point to a closed container.

In discussing this experiment, Moll and Tomasello say:

‘to understand pointing, the subject needs to understand more than the individual goal-directed behaviour. She needs to understand that ... the other attempts to communicate to her ... and ... the communicative intention behind the gesture’  
(Moll & Tomasello 2007)

Of course I don’t want to question this assertion. But I do want to suggest that in the context of joint action there is a way to respond reliably to informative pointing without understanding pointing at all. For if one knows that one is engaged in joint action with the person producing the point, one already knows what the (long-term) goal of the pointing action. The goal of the pointing is my goal, which is to find the reward. So in the context of a joint action, it should be no harder to understand the point than it is to understand the failed reach. Both are attempts to get the reward.

The pointing action, unlike the failed reach, is an **opaque means** of getting the object. But in the context of joint action this doesn’t matter because the your-goal-is-my-goal tells you that the goal of the point is to get the object.



The ‘my goal is your goal’ inference enables you to treat pointing as having the same goal as the failed reach. This amounts to *misunderstanding* pointing, of course. (The communicators’ goal is unlikely to be your goal.) But the misunderstanding is fruitful in the sense that it enables you to respond appropriately to the pointing, to make use of it.

So it is possible that the combination of minimal mindreading abilities with abilities to share goals is sufficient for understanding pointing actions in the context of joint actions.

As Ulf Liszkowski’s has demonstrated in a series of experiments, humans are unlike chimpanzees in they can understand and produce communicative actions involving pointing to inform (\*refs). In fact human children’s early abilities to understand and produce pointing gestures appear early in the second year of life. What I’m suggesting is that the emergence of these abilities might be facilitated by joint action. For it is possible to understand pointing gestures without *already* understand communicative intent.

**The idea is that in the context of joint action, communicative actions can be fruitfully misunderstood as ordinary goal-directed actions. But once an action has been given a function in joint action, it can be used to serve that function outside joint action contexts. And so it becomes genuinely communicative.**

Note that I am not suggesting that young children might fail to understand communicative intent. I am suggesting that they might first understand the goals of pointing actions without understanding communicative intent. But of course once they understand the goals of pointing actions within the context of joint action, it’s likely that they will be able to understand the goals of pointing actions outside the context of joint action too.<sup>14</sup>

Joint action may explain how individuals starting with a simple understanding of goals end up understanding communicative intentions.

This is one illustration of how capacities for joint action, even very simple forms of joint action, might be relevant to explaining the development or evolution of richer forms of cultural cognition.

---

<sup>14</sup> Contrast Csibra’s ‘two stances’ idea. The referential action understanding involves a “stance” (p. 455); teleological and referential action interpretation “rely on different kinds of action understanding” (? , p. 456); they are initially two distinct ‘action interpretation systems’ (although of course they come together later in development) (? , p. 456). In relation to Csibra’s, my suggestion is not that there is no referential stance. It’s rather that the referential stance might emerge from what he calls the ‘teleological stance’ together with abilities to engage in the sort of joint actions that are characterised by shared goals.

## 14. Conclusion

In conclusion I have suggested that there are forms of joint action which require only limited social cognition and which may play a role in the emergence of more sophisticated forms of social cognition, in development or evolution (or both).

[See last slide with diagram] On emergence, the idea was that abilities to engage in joint action combined with minimal social cognition enable humans to break into the Gricean circle and understand communicative intention. This in turn is one of the foundations on which abilities to communicate by language are built, and there is evidence that abilities to communicate by language in turn play a role in the emergence of full-blown mindreading abilities (\*refs). So this may be one route by which abilities to engage in joint action plus limited social cognition plays a role in the emergence of sophisticated forms of social cognition such as cognition of belief and other propositional attitudes.