

Mindreading and Joint Action: Philosophical Tools Notes for Lecture 7 What Is Shared Agency?

Stephen A. Butterfill
<ButterfillS@ceu.hu>

November 13, 2012

1. *Sources

- Slides and handout are based on one of the lectures from the Warwick 2011-12 *Joint Action and the Emergence of Mindreading* lectures (I also used a couple from the Milan talk on *Shared and Collective Goals*)
- paper: Shared and Collective Goals (*What Is Joint Action?* mind submission)
- talk notes: Tuebingen talk (*Which joint actions ground social cognition?*) (written up)
- paper: Shared Agency and Motor Representation (Budapest CEU version)

2. Introduction

What is shared agency? Although shared agency is a feature of many everyday events, it isn't easy to say what it is. One way to get an intuitive, pre-theoretical fix on phenomena of shared agency is to look at ways in which an exercise of shared agency can contrast with superficially similar exercises of individual agency in parallel.

When members of a flash mob in the Central Cafe respond to a pre-arranged cue by noisily opening their newspapers, they exercise shared agency. But when someone not part of the mob just happens to noisily

open her newspaper in response to the same cue, her action does not involve shared agency.¹ To give another example, two members of the mob exercise shared agency when they later walk to the metro station together as friends. But two people who merely happened to be walking to the metro station side by side would not be exercising shared agency (Gilbert 1990). These and other contrast cases invite the question, **How do activities involving shared agency differ from activities involving parallel agency only?**

2.1. Why it matters [**skip—this breaks the flow*]

Before taking this question head-on, it's helpful to constrain what would count as a good answer by stepping back and asking, Why is it useful to distinguish systematically between shared and individual agency?

Shared agency raises a tangle of scientific and philosophical questions. Psychologically we want to know which mechanisms make it possible (Sebanz et al. 2006; Vesper et al. 2010). Developmentally we want to know when shared agency emerges, what it presupposes and whether it might somehow facilitate socio-cognitive, pragmatic or symbolic development (Moll & Tomasello 2007; Hughes & Leekam 2004; Brownell et al. 2006). Phenomenologically we want to characterise what (if anything) is special about experiences of action and agency when shared agency is involved (Pacherie 2010). Metaphysically we want to know what kinds of entities and structures are implied by the existence of shared agency (Gilbert 1992; Searle 1994). And normatively we want to know what kinds of commitments (if any) are entailed by shared agency and how these commitments arise (Roth 2004). A philosophical account of shared agency that made it possible to systematically distinguish shared from individual agency may support investigation of these questions (as Bratman 2009 suggests).

2.2. Back to the Question

Let us return to the contrast cases.

How do cases involving shared agency differ from cases involving parallel agency only?

The first example shows that the difference between shared agency and parallel individual agency can't just be that the resulting actions have a common effect because merely parallel actions can have common effects too. And the second example shows that the difference can't just be a matter of coordination, because people who are merely happen to be walking side by

¹ See Searle (1990); in his example park visitors simultaneously run to a shelter, in once case as part of dancing together and in another case because of a storm. Compare Pears (1971) who uses contrast cases to argue that whether something is an ordinary, individual action depends on its antecedents.

side each other also need to coordinate their actions in order to avoid colliding. Note also that in both cases each individual's walking is intentional, so our intentionally walking together cannot be only a matter of our each intentionally walking.²

Perhaps, then, a notion of shared intention is needed to distinguish the two cases. Perhaps it is our acting on a shared intention that we walk together which distinguishes us from two strangers who happen to be walking side by side.

This view is common among philosophers psychologists:

'I take a collective action to involve a collective intention'
(Gilbert 2006, p. 5).

'the key property of joint action lies in its internal component [...] in the participants' having a "collective" or "shared" intention' (Alonso 2009, pp. 444-5).

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

'Shared intentionality is the foundation upon which joint action is built' (Carpenter 2009, 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' (Call 2009, p. 369).

But what could shared intention be?

3. Shared Intention

What is shared intention? Here is much disagreement. Some hold that it differs from ordinary intention with respect to the attitude involved (Searle 1990). Others have explored the notion that it differs from ordinary intention with respect to its subject, which is plural (Gilbert 1992; Helm 2008), or that it differs from ordinary intention in the way it arises, namely through team reasoning (Gold & Sugden 2007), or that it involves distinctive obligations or commitments to others (Gilbert 1992; Roth 2004). Opposing all such views, Bratman (1992, 2009) argues that shared intention can be realised by multiple

² This use of contrast cases resembles Pears (1971): he uses contrast cases to argue that whether something is an ordinary, individual action depends on its antecedents.

ordinary individual intentions and other attitudes whose contents interlock in a distinctive way.

To avoid having to take sides on what shared intention is, let us abstract the details and focus on features that everyone should agree on. Minimally, a shared intention stands to an activity involving shared agency in roughly the way that an ordinary, individual intention stands to an ordinary, individual action.

Shared intention stands to action involving shared agency as an ordinary intention stands to an ordinary intentional action.

Let us develop this parallel. Concerning ordinary individual agency we can ask,

What is the relation between a purposive action and the goal or goals to which it is directed?

As we saw last week, the notion of intention provides one way to answer this question. For an ordinary, individual intention represents an outcome, coordinates an agent's actions, and coordinates the agent's actions in such a way that, normally, this coordination would facilitate the occurrence of the represented outcome.

Concerning shared agency there is a parallel question:

What is the relation between a purposive joint action and the outcome or outcomes to which it is directed?

Here and onwards I'm going to use the term *joint action* to mean an exercises of shared agency. (This is just a stipulation; others use the term differently (e.g. Ludwig 2007).)

Just as we answered the original question by appeal to intention, so we can answer this question about joint action by appeal to shared intention. A shared intention involves there being a single outcome which each agent represents, where these representations coordinate the several agents' actions and coordinate them in such a way that, normally, the coordination would facilitate the occurrence of the represented outcome.

There is another feature of shared intention that should be uncontroversial on any account of it. Shared intentions or their components feature in practical reasoning alongside ordinary, individual intentions, and there are normative requirements which apply to combinations of individual and shared intentions. To illustrate, tonight there is a party and a ceremony. It is impossible for anyone to attend both, and this is common knowledge among us. We have a shared intention that we attend the ceremony together. While having this shared intention, I also intend to go to the party. Give our common knowledge, this combination of shared and individual intentions is irrational. Its irrationality is related to that which would be involved in my

individually intending to attend the ceremony while also intending to go to the party. In short, shared intentions or their components are inferentially and normatively integrated with ordinary, individual intentions.

It's important to see that we have been **moving in a circle**. The question was, What is shared agency? We're following most philosophers in supposing that we can answer this question by appeal to shared intention. That is, an activity involves shared agency when it involves actions that stand in an appropriate relation to a shared intention. Then we asked what shared intention is. And the answer is, it's something that stands in an appropriate relation to activities involving shared agency. I don't think all circles are bad, and this circle has provided us with some constraints on shared intentions could be. But it's important to see that, by moving in a circle, we have not said positively what shared intention is. For all we have said so far, it might be a mistake to assume that there is anything which stands to joint action as ordinary, individual intention stands to ordinary, individual action.

We need an account of shared intention.

4. Bratman on Shared Intention

Although there are many accounts of shared intention, the best developed account, to which no one has yet pushed a counterexample or decisive objection, is Michael Bratman's. ***

- '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' (Bratman 1993, p. View 4)

***Explain why the second condition is necessary: mafia cases on the way to New York.

5. A challenge

Is Bratman's account of shared intention and shared agency adequate to our needs?

One problem is that shared intentions are cognitively and conceptually demanding, as Guenther and Natalie note:

‘the contribution of lower-level processes to social interaction has hardly been considered. This has led philosophers to postulate complex intentional structures that often seem to be beyond human cognitive ability in real-time social interactions.’ (Knoblich & Sebanz 2008, p. 2022)

This might seem more obvious than it really is. On the substantial account given by Bratman, sharing intentions requires intentions about intentions (see Condition 2 in the quote above). 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, p. 332)

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.

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 joint action seems to impose cognitive demands and require conceptual sophistication close to the limits, or just beyond, of what humans are possible.

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.

Elsewhere Bratman says:

‘shared intentional agency consists, at bottom, in interconnected planning agency of the participants.’³

³ Bratman (2011, p. 11). See also Bratman (2011, p. 5): ‘We begin with planning agents.’

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.⁴

So Guenther and Natalie are right insofar as the account is conceptually and cognitively demanding.

5.1. A conjecture

But so what if, on Bratman's account, shared intention requires conceptual and cognitive sophistication? Maybe this is just how it is with joint action. So I don't think there's an objection to the account here yet. There's just a question. The question is whether there are forms of shared agency which are not demanding in the way that Bratman's account suggests.

There's a further reason why we might want a less demanding kind of shared agency ...

Several researchers have conjectured that shared agency in some sense grounds sophisticated forms of cognition. In this vein, Sebanz and Knoblich conjecture that:

'[F]unctions traditionally considered hallmarks of individual cognition originated through the need to interact with others ... perception, action, and cognition are grounded in social interaction.' (Knoblich & Sebanz 2006, p. 103)

Relatedly, Moll and Tomasello offer a 'Vygotskian Intelligence Hypothesis':

'the unique aspects of human cognition ... were driven by, or even constituted by, social co-operation.' (Moll & Tomasello 2007, p. 1)

And Sinigaglia and Sparaci write:

'human cognitive abilities ... [are] built upon social interaction' (*ref Sinigaglia and Sparaci 2008)

Although these researchers write about 'cultural interactions' and 'social interactions', I think it's clear from the contexts that their target is shared agency.

These bold ideas are the inspiration for a more limited conjecture.

Conjecture The prior existence of capacities for shared agency partially explains how sophisticated forms of mindreading emerge in evolution or development (or both).

⁴ This claim is argued for in detail in ?

This conjecture faces several objections. First, don't don't sophisticated forms of mindreading emerge before shared agency? We've already dealt with this objection by distinguishing minimal from full-blown mindreading in earlier lectures. The discussion from those lectures does not show that the objection is actually mistaken; but it does show that whether the objection holds good or not depends on further empirical research.

The second objection is relevant now. It is this:

1. All shared agency involves shared intention.
2. Shared intention requires sophisticated mindreading.

Therefore:

3. The prior existence of capacities for shared agency could play no significant role in explaining how sophisticated forms of mindreading emerge.

The argument does not show that there is no explanatory role for capacities for shared agency at all. It may be that the benefits shared agency brings explain *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 mindreading by appeal to shared agency involving shared intention. But what we can't do is explain *how* we acquire mindreading by appeal to anything involving shared intention. Shared intention presupposes, and so cannot explain, sophisticated mindreading.

In response to this argument some philosophers have attempted to reject the second premise (e.g. Tollefsen 2005; Pacherie 2012). This would mean rejecting Bratman's account of shared intention, and formulating and defending an account. This would be very difficult because Bratman's account has proved quite resistant to objections.

Instead of taking this route, I want to consider rejecting the first premise, (1) above.

Is it reasonable to reject the first premise? You might think not, for this reason: As philosophers, the topic is not just doing things together but intentionally doing them together.⁵ But even if we accept this, I think we shouldn't assume that all shared agency involves shared intention. After all, the parallel with ordinary, individual agency suggests otherwise. [*More on this in Budapest version of *Shared Agency and Motor Representation* paper.]

⁵ I.e. here we argue about *intentional joint action*. ***

6. Shared Agency without Shared Intention: First attempt

Our question is whether we can defend the Conjecture by identifying a kind of shared agency that does not involve shared intention.

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 (Ludwig 2007, p. 366).

7. *Missing bits

See the *What Is Joint Action? A Modestly Deflationary Account (Shared and Collective Goals)* paper.

7.1. 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 most wants and expects each of the other agents to perform activities directed to the goal; and
 - (b) each agent most wants and 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.

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.

8. Conclusion

These three notions—shared goal, collective goal and distributive goal—identify three ways in which a joint action could be related to its goal. These do not require sophisticated mindreading; in fact they require only goal ascription. They provide building blocks for characterising shared agency without shared intention.

Why should we care whether there could be shared agency without shared intention? One motive is the Conjecture that the prior existence of capacities for shared agency may partially explain how sophisticated forms of mindreading emerge in evolution or development (or both).

Because having a shared intention would require sophisticated mindreading—mindreading close to the limits of what human adults are capable of—, to show that the Conjecture is coherent we need to establish that there can be shared agency without shared intention. Forms of joint action which involve shared intention *presuppose* sophisticated social cognition and so cannot *explain how* it emerges in either evolution or development.

This is why the notions of distributive, collective and shared goals are interesting. They provide the basis for forms of shared agency which require only limited mindreading abilities. 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, or that fully understanding shared agency involves understanding shared intention too. 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 the roles

of shared agency in evolution or development (or both), we probably need to start with the simplest kinds of shared agency.

References

- Alonso, F. M. (2009). Shared intention, reliance, and interpersonal obligations. *Ethics*, 119(3), 444–475.
- Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review*, 101(2), 327–341.
- Bratman, M. (1993). Shared intention. *Ethics*, 104, 97–113.
- Bratman, M. (2009). Modest sociality and the distinctiveness of intention. *Philosophical Studies*, 144(1), 149–165.
- Bratman, M. (2011). Acting over time, acting together (draft).
- Brownell, C. A., Ramani, G. B., & Zerwas, S. (2006). Becoming a social partner with peers: cooperation and social understanding in one- and two-year-olds. *Child Development*, 77(4), 803–21.
- Call, J. (2009). Contrasting the social cognition of humans and nonhuman apes: The shared intentionality hypothesis. *Topics in Cognitive Science*, 1(2), 368–379.
- Carpenter, M. (2009). Just how joint is joint action in infancy? *Topics in Cognitive Science*, 1(2), 380–392.
- Gilbert, M. (1992). *On Social Facts*. Princeton, NJ: Princeton University Press.
- Gilbert, M. (2006). Rationality in collective action. *Philosophy of the Social Sciences*, 36(1), 3–17.
- Gilbert, M. P. (1990). Walking together: A paradigmatic social phenomenon. *Midwest Studies in Philosophy*, 15, 1–14.
- Gold, N. & Sugden, R. (2007). Collective intentions and team agency. *Journal of Philosophy*, 104(3), 109–137.
- Helm, B. W. (2008). Plural agents. *Nous*, 42(1), 17–49.
- Hughes, C. & Leekam, S. (2004). What are the links between theory of mind and social relations? review, reflections and new directions for studies of typical and atypical development. *Social Development*, 13(4), 590–619.
- Knoblich, G. & Sebanz, N. (2006). The social nature of perception and action. *Current Directions in Psychological Science*, 15(3), 99–104.

- Knoblich, G. & Sebanz, N. (2008). Evolving intentions for social interaction: from entrainment to joint action. *Philosophical Transactions of the Royal Society B*, 363, 2021–2031.
- Ludwig, K. (2007). Collective intentional behavior from the standpoint of semantics. *Nous*, 41(3), 355–393.
- Moll, H. & Tomasello, M. (2007). Cooperation and human cognition: the vygotkian intelligence hypothesis. *Philosophical Transactions of the Royal Society B*, 362(1480), 639–648.
- Pacherie, E. (2012). Intentional joint agency: Shared intention lite. *Synthese*, forthcoming.
- Pacherie, E. (forthcoming 2010). The phenomenology of joint action: Self-agency vs. joint-agency. In A. Seeman (Ed.), *Joint Action*. MIT Press.
- Pears, D. (1971). Two problems about reasons for actions. In A. M. R. Binkley, R. Branaugh (Ed.), *Agent, Action and Reason* (pp. 128–153). Oxford: Oxford University Press.
- Roth, A. S. (2004). Shared agency and contralateral commitments. *The Philosophical Review*, 113(3), 359–410.
- Searle, J. R. (1990). Collective intentions and actions. In P. Cohen, J. Morgan, & M. Pollack (Eds.), *Intentions in Communication* (pp. 90–105). Cambridge: Cambridge University Press. Reprinted in Searle, J. R. (2002) *Consciousness and Language*. Cambridge: Cambridge University Press (pp. 90–105).
- Searle, J. R. (1994). *The Construction of Social Reality*. New York: The Free Press.
- Sebanz, N., Bekkering, H., & Knoblich, G. (2006). Joint action: Bodies and mind moving together. *Trends in Cognitive Sciences*, 10(2), 70–76.
- Tollefsen, D. (2005). Let's pretend: Children and joint action. *Philosophy of the Social Sciences*, 35(75), 74–97.
- Vesper, C., Butterfill, S., Knoblich, G., & Sebanz, N. (2010). A minimal architecture for joint action. *Neural Networks*, 23(8-9), 998–1003.