

# Joint Action and Knowing Others' Minds: How social interaction grounds mindreading

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

November 10, 2011

## Abstract

\*\*\*

## 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.

## 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>

---

<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 (??).

<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).

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

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. 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 participants 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 ...

### 3.1. 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>3</sup>

---

<sup>3</sup> The interdependent roles of beliefs, desires and goals in producing action mean there

### 3.2. 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>4</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.

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,

---

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>4</sup> \*What notion of joint action is needed here? Any will do as long as it involves distributive goals.

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.

### 3.3. 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>5</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.

## 4. 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

---

<sup>5</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.

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.

## References

- Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adults' performance on a non-inferential theory of mind task. *Cognition*, 106, 1093–1108.
- Apperly, I. A. & Butterfill, S. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, 2009(116), 4.
- Apperly, I. A., Samson, D., & W., H. G. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology*, 45(1), 190–201.
- Astington, J. & Baird, J. A. (Eds.). (2005). *Why Language Matters for Theory of Mind*. Oxford: Oxford University Press.
- Butterfill, S. & Apperly, I. A. (submitted 2011). How to construct a minimal theory of mind. [http://butterfill.com/papers/minimal\\_theory\\_of\\_mind/](http://butterfill.com/papers/minimal_theory_of_mind/).
- Clements, W., Rustin, C., & McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: a training study of false belief. *Developmental Science*, 3(1), 81–92.
- Csibra, G. (2003). Teleological and referential understanding of action in infancy. *Philosophical Transactions: Biological Sciences*, 358(1431), 447–458.
- Hare, B. & Tomasello, M. (2004). Chimpanzees are more skilful in competitive than in cooperative cognitive tasks. *Animal Behaviour*, 68(3), 571–581.



- 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.
- Leslie, A. (2005). Developmental parallels in understanding minds and bodies. *Trends in Cognitive Sciences*, 9(10), 459–62.
- McKinnon, M. C. & Moscovitch, M. (2007). Domain-general contributions to social reasoning: Theory of mind and deontic reasoning re-explored. *Cognition*, 102(2), 179–218.
- Moll, H. & Tomasello, M. (2007). Cooperation and human cognition: the vygotskian intelligence hypothesis. *Philosophical Transactions of the Royal Society B*, 362(1480), 639–648.
- Perner, J. & Lang, B. (1999). Development of theory of mind and executive control. *Trends in Cognitive Sciences*, 3(9), 337–344.
- Sabbagh, M. (2006). Executive functioning and preschoolers' understanding of false beliefs, false photographs, and false signs. *Child Development*, 77(4), 1034–1049.
- Slaughter, V. & Gopnik, A. (1996). Conceptual coherence in the child's theory of mind: Training children to understand belief. *Child Development*, 67, 2967–2988.
- Wellman, H., Cross, D., & Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false-belief. *Child Development*, 72(3), 655–684.
- Wimmer, H. & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13, 103–128.