

Origins of Mind: Lecture Notes

Lecture 05

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

October 31, 2013

title-slide

Knowledge of Mind

slide-3

The challenge is to explain the emergence, in evolution or development, of mindreading. Let me explain.

slide-6

Mindreading is the process of identifying mental states and purposive actions as the mental states and purposive actions of a particular subject.

slide-7

Researchers sometimes use the term 'theory of mind'. 'In saying that an individual has a theory of mind, we mean that the individual imputes mental states to himself and to others' (Premack and Woodruff 1978, p. 515)

slide-10

So, to be clear about the terminology, to have a theory of mind is just to be able to mindread, that is, to identify mental states and purposive actions as the mental states and purposive actions of a particular subject.

slide-14

So the challenge is to explain the emergence of mindreading. You know (let's say) that Ayesha believes Beatrice is in the library. Humans are not born knowing individuating facts about others' beliefs. How do they come to be in a position to know such facts? Meeting this challenge initially seems simple. But, as you'll see, we quickly end up with a puzzle. I think this puzzle requires us to rethink what is involved in having a conception of the mental.

slide-15

I shall focus on awareness of others' beliefs to the exclusion of other mental states. There's no theoretical reason for this; it's just a practical thing. And what we learn about belief will generalise to other mental states.

maxi_story

How can we test whether someone is able to ascribe beliefs to others? Here is one quite famous way to test this, perhaps some of you are even aware of it already. Let's suppose I am the experimenter and you are the subjects. First I tell you a story ...

slide-20

Here's the really surprising thing. Children do really badly on this until they are around four years of age. And they seem to develop the ability to pass this task only gradually, over months or years. (There's something else that isn't surprising to most people but should be: adult humans not only nearly always provide the answer we're calling 'correct': they also believe that there is an obviously correct answer and that it would be a mistake to give any other answer. I'll return to this point later.) (NB: The figure is not Wimmer & Perner's but drawn from their data.)

slide-21

There's been some stuff in the press recently about bad science, mainly some dodgy methods and failures to replicate.

slide-22

So you'll be pleased to know that a meta-study of 178 papers confirmed Wimmer & Perner's findings. Now there is clearly some variation here. That's because different researchers implemented different versions of the original task. We can use the meta-analysis of these experiments as a shortcut to finding out what sorts of factors affect children's performance.

slide-23

One factor that seems to make hardly any difference is whether you ask children about others' beliefs or their own beliefs. To repeat, you get essentially the same results whether you ask children about others' beliefs or their own beliefs. Children literally do not know their own minds.

slide-24

What happens if we involve the child by having her interact with the protagonist? The task becomes easier for children of all ages, but the transition is essentially the same (participation does not interact with age Wellman et al. 2001, pp. 665-7).

slide-25

Finally, although there are some cultural differences, you get the same transition in seven different countries.

slide-26

So our challenge was to explain the emergence of mindreading. At this point, up until around, it seemed quite straightforward to most researchers. We seemed to know that children are unaware of mental states until around four years. And a lot of studies looked at which factors affect their acquiring this awareness. These studies showed that executive function, language and rich forms of social interaction are all important. All of this supported something like the story that Sellars tells in his famous Myth of Jones.

slide-27

todo: describe Sellars' myth

slide-28

But there was a big surprise in store for us.

unit_411

Mindreading: First Puzzle

slide-30

There is a puzzle about when humans can first know individuating facts about others' beliefs. To understand the origins of this knowledge we need to understand the puzzle. So I'm going to reveal the puzzle to you. But let me start with a bit of background.

slide-31

Recall the experiment that got us started. These experimenters added an anticipation prompt and measured to which box subjects looked first (Clements and Perner 1994). (Actually they didn't use this story; theirs was about a mouse called Sam and some cheese, but the differences needn't concern us.)

slide-36

What got me hooked philosophical psychology, and on philosophical issues in the development of mindreading in particular was a brilliant finding by Wendy Clements who was Josef Perner's phd student. These findings were carefully confirmed (Clements et al. 2000; Garnham and Ruffman 2001; Ruffman et al. 2001). Around 2000 there were a variety of findings pointing in the direction of a conflict between different measures. These included studies on word learning (Carpenter et al. 2002; Happé and Loth 2002) and false denials (Polak and Harris 1999). But relatively few people were interested until ...

slide-88

The challenge is to explain the emergence, in evolution or development, of mindreading. Initially it looked like this was going to be relatively straightforward and involve just language, social interaction and executive function. So a Myth of Jones style story seemed viable. But the findings of competence in infants of around one year of age changes this. These findings tell us that not all abilities to represent others' mental states can depend on things like language. And, as I've been stressing, these findings also create a puzzle. The puzzle is, roughly, how to reconcile infants' competence with three-year-olds' failure.

slide-89

The puzzle is a little bit like the puzzle we had in the case of knowledge of physical objects. But it's also different. In the case of physical objects, the conflict was between measures involving looking and measures involving searching. In this case it's different, because on the infant side there is not just looking but also acting (e.g. helping) and even communicating.

slide-90

Can we solve the puzzle by appeal to core knowledge (or 'modularity')? The difference in measures is a hopeful sign that we can. But the fact that representations of others' minds influence 1-year-olds' actions (e.g. in communicating and helping) complicates things because we imagine modules as inferentially isolated from practical reasoning. Looking at a further puzzle will help us.

unit_421

Mindreading: Second Puzzle

slide-109

So how does this second puzzle bear on our overall objective? The challenge is to explain the emergence, in evolution or development, of mindreading. The fact that mindreading is sometimes automatic and sometimes not is good evidence that mindreading is not a single thing. Rather there are multiple kinds of process or system involved in mindreading. From this we can conclude two things. First, in

this respect, the cases of mindreading is much like the cases of colour and physical objects. These also involved multiple systems; this took the form of a distinction between core knowledge and knowledge proper. Second, we have further grounds for thinking that some forms of mindreading are modular.

slide-110

[These slides are from a joint presentation with Ian Apperly (BCCCD)] What is the relation between infant and adult capacities for mindreading? It's important at this point to recognise that there are two quite different possibilities.

unit_431

Modules and Cognitive Efficiency

slide-117

We can resolve both puzzles by appeal to the idea that there are modular & non-modular mindreading processes. (Cf. physical objects).

slide-118

Explain how this works in the case of each puzzle. *todo* modify the slide to illustrate the solution

slide-122

First, what do I mean by cognitive efficiency. Some tasks require mental effort. For example, suppose I ask you to count from one to one hundred omitting each prime number. This is not the kind of thing most people can do while washing up. It requires attention, inhibitory control and working memory. By contrast, many adults can count in the ordinary way from one to one hundred while washing up. Counting has become routine, habitual.

slide-125

So much for cognitive efficiency. Now why suppose that modularity (or core knowledge) requires cognitive efficiency? To invoke modularity, we need to understand how mindreading could be (i) automatic and (ii) present in pre-linguistic infants with limited working memory & executive function; both (i) and (ii) mean we need to understand how it could be cognitively efficient. (The situation is a bit like this: we want to say you can perceive others mental states; but on the face of it, mental states are exactly the sort of things that are not available to perception.)

slide-126

We saw earlier that mindreading in four year olds and adults is cognitively demanding. And there's good reason to think that it should be. If anything should consume scarce cognitive resources ... Now appeal to modularity doesn't explain how mindreading might somehow be efficient. Suppose someone could find prime factors incredibly quickly. It wouldn't be a satisfying explanation to just say that she had a module for finding prime factors. We'd also need an algorithm that her module could be implementing consistently with her performance. So (a) efficiency points to modularity but (b) efficiency requires explanation and (c) gesturing at modularity doesn't explain efficiency. To see how mindreading could be cognitively efficient, we need to reject a dogma.

unit_44I

Minimal Theory of Mind

unit_45I

Signature Limits Generate Predictions

slide-169

So let me conclude. The challenge we have been addressing was to understand the emergence of mindreading. Initially this seemed straightforward: you learn this from social interaction using language as a tool (compare Gopnik's theory theory). However, the discovery that abilities to track beliefs exist in infants from around 7 months or earlier initially suggested a different picture: one on which mindreading

was likely to involve core knowledge. But, as always, things are not so straightforward. The evidence is apparently conflicting. There are actually two conflicts, not one: developmentally (A- & B-tasks) and in adults, automaticity. The existence of two puzzles gives us confidence that the conflict is not merely a methodological artefact. The solution is to recognise that there are modules, but there was an obstacle to the hypothesis that mindreading could be modular (*cognitive efficiency) In constructing minimal theory of mind we've earned the right to solve them by appeal to modularity. (NB: it's modularity rather than mTm that explains the discrepancy; mTm is important because (i) it explains efficiency; and (ii) it generates predictions via signature limits) Now the idea that there are both modular and non-modular mindreading enables us to solve the two puzzles (developmentally (A- & B-tasks) and in adults, automaticity). However, this resolution of the puzzles doesn't answer our overall challenge about the origins of knowledge of other minds. In fact it complicates the account of the origins of knowledge of other minds, makes the challenge seem harder rather than easier to meet. We can't give a theory theory / learning account; and we also can't give a straightforward core knowledge account. Instead we need something a bit more complex. I haven't tried to offer an account of what that thing is yet, and that isn't the point of this lecture. But let me close by describing how I would approach it. A key issue is the relation between infant and adult competence ...

slide-170

Infant competence retained in adults (matching signature limits from Low & Watts), as in the cases of colour (and probably physical objects too). Now if we accept this, it is tempting to conjecture that the later developing mindreading abilities involve a process of rediscovery ...

slide-171

- theme A: explain the origins of knowledge of others minds : development as re-discovery. There is a modular capacity (= core knowledge). But this doesn't lead to adult-like understanding for years, and the acquisition of adult-like understanding hinges on language; may involve completely different model of mental states.

References

- Carpenter, M., Call, J., and Tomasello, M. (2002). A new false belief test for 36-month-olds. *British Journal of Developmental Psychology*, 20:393-420.
- Clements, W. and Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, 9:377-395.
- Clements, W., Rustin, C., and McCallum, S. (2000). Promoting the transition from implicit to explicit understanding: a training study of false belief. *Developmental Science*, 3(1):81-92.
- Garnham, W. and Ruffman, T. (2001). Doesn't see, doesn't know: is anticipatory looking really related to understanding or belief. *Developmental Science*, 4(1):94-100.
- Happé, F. and Loth, E. (2002). 'theory of mind' and tracking speakers' intentions. *Mind and Language*, 17(1-2):24-36.
- Polak, A. and Harris, P. (1999). Deception by young children following noncompliance. *Developmental Psychology*, 35(2):561-568.
- Premack, D. and Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences*, 1(04):515-526.
- Ruffman, T., Garnham, W., Import, A., and Connolly, D. (2001). Does eye gaze indicate implicit knowledge of false belief? charting transitions in knowledge. *Journal of Experimental Child Psychology*, 80:201-224.
- Wellman, H., Cross, D., and Watson, J. (2001). Meta-analysis of theory of mind development: The truth about false-belief. *Child Development*, 72(3):655-684.