Intentionalism and Mindreading
Intention & Acts of Meaning Seminar, Week 5

Daniel W. Harris

March 3, 2016

1 Questions

- When do children develop their capacity (or capacities) for mindreading?

- To what extent do adults communicate using mindreading in the way that intentionalism entails?

- What are the implications of the current psychological literature for intentionalist theories of communication and utterance-type meaning?

- What are intentionalists' (controversial) commitments about mindreading?

2 False-Belief Tasks

- Massively influential methodology starting with Wimmer and Perner (1983).

- Attempt to answer the question: Can S represent the beliefs of a target agent in a way that distinguishes S's beliefs from their own?

- Operationalized by measuring behavior that indicates that S represents the target agent a belief that S knows to be false.

2.1 1983–2005: Tests for Explicit False Beliefs

For two decades, it was orthodox to think that children start passing false-belief tasks when they are about 4. This was taken as evidence about when the human mindreading capacity begins to develop. There were *many* studies, but the classic is by Wimmer & Perner (1983):

> …subjects observed how a protagonist put an object into a location x and then witnessed that in the absence of the protagonist the object was transferred from x to location v. Since this transfer came as a surprise they had to assume that the protagonist still believed that the object was in x. Subjects had to indicate where the protagonist will look for the object at his return. None of the 3–4-year old, 57% of 4–6-year old, and 86% of 6–9-year old children pointed correctly to location x in both sketches. (Wimmer and Perner, 1983, 103–104)

It is important that by saying that subjects "had to indicate" a location, Wimmer & Perner mean that the subjects were explicitly asked one of two questions, and had to answer by pointing or speaking (Wimmer and Perner, 1983, 109):

BELIEF QUESTION: 'Where will Maxi look for the chocolate?'

UTTERANCE QUESTION: 'Where will Maxi say the chocolate is?'

Nearly all tasks until 2005 were **explicit** in this sense. For a meta-analysis of these studies, seeWellman et al. (2001).

2.2   The Pre-2005 Dilemma for Intentionalism

As of the early 2000's, it seemed clear to most that 3-year-olds couldn't mindread. (The most prescient dissenter from this consensus was Alan Leslie.)

So intentionalists[1] faced the following objection (formulated by Santana (MS), following Breheny (2006):[2]):

1. 3-year-olds are competent language users.

2. 3-year-olds lack theory of mind.

3. Therefore, theory of mind is not required for linguistic competence.

What responses were available?

Deny 1: 3-Year Olds Aren't Competent Language Users

- Maybe 3-year-olds are doing something that resembles full-blown language use, but not engaging in ostensive-inferential communication?

Some evidence: Joseph and Tager-Flusberg (2010) found that 3-year-olds don't take the audience's state of mind into account when using definite descriptions (they use DDs egocentrically). (Interestingly, Happé and Loth (2002) found the opposite result when using DDs containing made-up predicates, like 'sefo'.)

But there just isn't much evidence for this hypothesis.

Moreover, some evidence suggested that 3-year-olds could do indexicals but not implicature. This could be used to support the idea that "literal" communication uses non-Gricean mechanisms but non-literal communication is Gricean. That would be a bad conclusion for intentionalism.

---

[1] Many non-intentionalist theories of communication nonetheless also rely on elaborate claims about mindreading, including those of Lewis, Stalnaker, etc.

[2] It's worth noting that although Breheny's paper was published in 2006, it had been written at least a couple of years earlier, and he seems not to have known yet about the implicit data that had begun to emerge. Publication is slow.

Deny 2: 3-Year-Olds *Do* Can Mindread Despite Failing the Tests

- Posit a "comprehension module" that is specialized for ostensive-inferential communication, and that develops before our regular mindreading capabilities (the ones needed to pass a false belief task)(Sperber and Wilson, 2002).

Sperber & Wilson (2002) also offer some other reasons for thinking that ostensive mindreading is special. (See DeChant (MS) for a point-by-point rejoinder.)

But: This proposal strikes many as ad hoc, and it undermines some of the original attraction of intentionalism—i.e., that communication relies on an independently motivated psychological capacity for action explanation/interpretation.

- Argue that 3-year-olds have the conceptual capacity for mindreading, but that they lack some cognitive resources (such as sufficient executive function) for performance (Fodor, 1992; Kikuno and ans Fenja Ziegler, 2007; Leslie, 1994; Leslie et al., 2005; Roth and Leslie, 1998).

As we'll see, this is now a major position.

But as of 2004, it was supported by mainly indirect evidence. E.g., the fact that young children engage in pretense (Leslie, 1987), and that they show evidence of processing effort when performing false-belief tasks (Kikuno and ans Fenja Ziegler, 2007; Leslie et al., 2005).

And so this was not a mainstream view before 2005.

Weaken Intentionalism

- Maybe we can come up with less cognitively demanding formulations of intentionalism (Csibra, 2010; Moore, MS)?

These don't seem to simplify communication enough to do what is wanted. And when they do—e.g., by leaving the content of the informative intention out of the content of the communicative intention—they give up much of the explanatory power of intentionalism (Thompson, 2014, §3.1).

- Maybe intentionalism is just a rational reconstruction?

Okay, but how does it really work?

Note that calling intentionalism a rational reconstruction is a much bigger concession that calling Grice's theory of implicature a rational reconstruction. The latter is a much more plausible claim.

- Maybe intentionalism is an idealized model (Santana, MS; Thompson, 2008)?

Santana argues that idealized models can give us predictive and explanatory benefits in the absence of a fuller theory by isolating special cases of communication (when everything is going well) or by isolating core causal mechanisms (the ones at work in adult comprehension).

Maybe, but we would like to know more about what the relevant mechanisms are, and what's going on in the non-ideal cases. A worrying possibility is that the kids have the core mechanisms and adults have add-ons.

The difference between this and the rational-reconstruction idea is also tough to pin down.

2.3   2005–Present: Tests for Implicit False Beliefs

Onishi & Baillargeon (2005) devised the first *implicit* false-belief task that seemed to show that infants (15-month-olds) could mindread. Since then, many other experiments have shown the same thing, including with subjects as young as 10 months (Luo, 2011).

What makes these tests *implicit*? They infer false-belief attribution from subjects' behavior, rather than asking them to verbalize it or otherwise communicating it directly. There are several paradigms:

SURPRISED LOOKING: While S watches, A puts an object in box 1, leaves the room, then B comes in and moves object to box 2. When the original agent comes back, S stares longer if A looks for the object in Box 2. This staring is taken as a measure of surprise—A's behavior defied S's expectation that A would look for the object in Box 1. This is evidence that S attributed a false belief to A (Onishi and Baillargeon, 2005).

ANTICIPATORY LOOKING: Similar to last case, but uses an eye tracker to check which box infants look at *before* A chooses. Infants reliably look at the box

where A hid the object, not at the object's location. This is taken as evidence of the infant's expectation of A's behavior, which is evidence of false-belief attribution (Southgate et al., 2007).

ACTIVE HELPING: Similar to above, but now the boxes are locked when A returns and only the child knows how to open them. While A tries to open Box 1, the child is asked for help. Most children open Box 2 "apparently assuming that [A] was trying to open the first box in order to get the toy—but had a false belief as to where the toy was" (Buttelmann et al., 2009, 339).

3   Agency Detection

Another source of evidence for infant mindreading comes from goal-detection experiments, many of which are summarized by (Carey, 2009, ch.5).

Many of these use experimental paradigms similar to SURPRISE LOOKING, but in which subjects watch an animation (or a live action stage play). An example is Figure 5.1 (borrowed from (Carey, 2009, 160)).
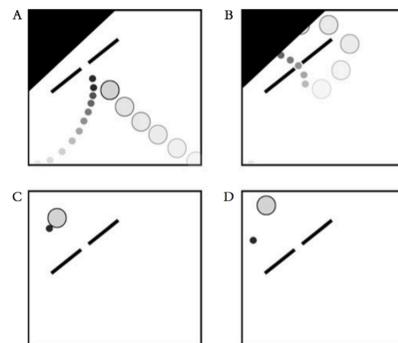

Figure 5.1. Schematic depiction of habituation trials in Csibra et al., 1999. Reprinted from Csibra, G., Gergely, G., Koos, O., & Brockbank, M. (1999). Goal attribution without agency cues: The perception of "pure reason" in infancy. *Cognition, 72*, 237–267, with permission from Elsevier. The small ball goes into motion, passing through the small gap in the barrier, then going out of sight. The large ball appears to follow it toward the gap, then goes around the barrier before passing out of sight.

These experiments find that infants (as young as 3-months-old!) interpret the

movements of animated figures by attributing goals to them. For example, after seeing several trials in which the big ball stops next to the small ball, they seem to attribute the goal of chasing the small ball to the big ball. In subsequent trials with different obstacles, they stare longer when this expectation isn't met.

There is also evidence that infants as young as 14 months attribute communicative intentions:

> In the context of a hiding game, an adult indicated for the child the location of a hidden toy by giving a communicative cue: either pointing or ostensive gazing toward the container containing the toy. To succeed in this task children had to do more than just follow the point or gaze to the target container. They also had to infer that the adult's behaviour was relevant to the situation at hand – she wanted to inform them that the toy was inside the container toward which she gestured. Children at all three ages successfully used both types of cues. (Behne et al., 2005, 492)

Finally, Carey and coauthors present evidence that infants can distinguish between an object with a goal and an object that is behaving according to the goals of another agent who is controlling the object (Carey, 2009, 167–168). They also show that infants can distinguish between the goals of different agents in a situation, and that they attribute intentional dispositions that persist in future trials (Carey, 2009, 169).

4   Implicit Mindreading Vs. Explicit Mindreading

What is the difference between implicit and explicit mindreading? Which do we require for communication?

(Of course, to say that there are two kinds of mindreading is a bit misleading. Really there are two ways of measuring it that return different results. The question is, why?)

Some Bad Ways of Explaining the Implicit/Explicit Distinction

Carruthers considers and rejects several simple answers:

- Implicit representations of mental states are "represented but not judged", whereas explicit representations are judged. (This is a tenuous distinction, and it fails to account for active-helping paradigms.)

- Implicit representations are unconscious, but explicit ones are conscious. (But why think that we need to be conscious of everything we assert? In other words: consciousness can't explain what changes at 4-years-old.)

- Implicit attitudes are embodied in the system but not represented in a structured way; explicit attitudes are represented (e.g., in a language of thought). (But merely embodied attitudes aren't available as premises in inference, whereas the infants' belief attributions are.)

Another possibility is that the implicit tests are measuring a completely different cognitive competence than mindreading.

- Perner and Ruffman (2005) and others have defended simple accounts of infant mindreading in terms of simple behavior rules like "People will search for an object where it was last in their line of sight".

- Butterfill and Apperly (2013) argue that infants could make the right predictions by attributing *functions* in place of intentions and *fields* (primitive relations to their non-occluded immediate environments) in place of beliefs or perceptual states. (For a similar argument, see Burge 2015.)

If they're right, then intentionalism is back where it was pre-2005.

But as more experiments are done, the more these explanations look ad hoc.

Moreover, it seems that they already can't explain many of the more complex results turned up in the labs of Carey and Tomasello. Carruthers argues roughly this (Carruthers, 2013, §2.5).

A Competence-Performance Distinction?

> The most plausible suggestion is that it is something about language production (or the production of communicative actions generally, including pointing gestures …that disrupts successful performance in verbal false-belief tasks (Scott and Baillargeon, 2009). In particular, it seems reasonable to think that the triple burden imposed on the mindreading system

in such tasks could prove too much for infants. For in order to succeed in such a task the child needs to do three things: (1) it has to process and keep in mind the mental states attributed to the target agent, as well as have these interact with its own planning systems to generate expectations for behavior; (2) it has to process the speech of the experimenter and figure out the underlying communicative intention; and (3) it has to formulate an action that would serve to communicate the target agent's mental states or likely actions to the experimenter. Since mindreading will be implicated in all three things (Sperber and Wilson, 2002), it makes sense that a capacity to pass verbal false-belief tasks might depend upon maturational expansion of the processing resources available to the mindreading faculty, or on increasing efficiency in the interactions between thought-attribution systems and executive systems, or both. (Carruthers, 2013, 153)

Does one need explicit mindreading for ostensive-inferential communication, or is implicit mindreading good enough?

Implicit is enough, I think. After all: we don't constantly narrate what we're doing while we're talking.

5   Egocentric Reasoning in Adults

Keysar (2007) argues that, like children failing the false-belief task, adults often (but don't always) reason "egocentrically" when communicating. For example, they choose a description of an object that fits the object from their point of view, but not from what they know about their hearer's point of view.

The effect is increased as a result of cognitive load:

We manipulated the external "cognitive load" by asking subjects to keep in mind either two (low load) or five (high load) sets of numbers while following instructions. Indeed, with a high external load subjects were much more egocentric than with low external load, and behaved like subjects who have a low working memory capacity. (Keysar, 2007, 75)

Does this threaten intentionalism? No. The fact that adults often don't reason egocentrically (particularly when they're not distracted) shows that intentionalism is the right theory of the good cases. (Children also communicate non-egocentrically in some experiments—see Happé and Loth (2002).)

But it does give us something that we need to explain. Why does ostensive-inferential communication exhibit these particular patterns of breakdown?

6   A Hypothesis

Maybe something about the mechanisms behind mindreading explains the typical patterns of breakdown in infants and adults under cognitive load?

6.0.1   Theory Theory

Mindreading is like scientific inference. We have a theory of how minds work, and we infer their mental states from their behavior together with the theory (Dennett, 1969, 1971, 1987; Sellars, 1956).

How would a theory theorist explain the fact that we reason egocentrically when under load?

6.0.2   Simulationism

An alternative to theory-theory is simulationism, according to which we model others' beliefs and plans by simulating them using our own belief and planning mechanisms.

Here's how Carruthers describes the process by which infants attribute goals to shapes in an animation:

What happens when an infant forms an expectation about an agent's behavior, then, will be something like this. On the basis of the previous familiarization trials, the core mindreading system ascribes to the agent the goal of being located next to the other agent (say). …this goal is passed along to the infant's own planning system in the form of the query, 'How does one get there [the position of the goal] from there [the position of the agent]?' The planning system sets to work to construct a plan, constrained

by observable features of the environment, and drawing as needed from among the infant's own currently accessible beliefs. The resulting plan ('Move in a straight line' or 'Move around that obstacle') is attributed to the target agent, thereby forming an expectation about what she should do (either prospective or retrospective). And it is this that is violated if the agent takes a circuitous route in the absence of an obstacle. (Carruthers, 2013, 144)

This theory is parsimonious, and it goes some way to explaining how mindreading can be so efficient. (It's less clear why it mindreading should be fast and automatic, however, given that planning often isn't.)

One problem: how does this theory explain why we don't always reason egocentrically? How do we reason from the perspective of the target agent—i.e., taking their mental states into account—instead of simply relying on our own beliefs?

Simulationists have to add an extra, "imaginative" step in order to account for this.

> Taking into account another's ignorance or false belief when predicting or explaining their behavior requires imaginative modifications of one's own beliefs, according to the simulation theory. Thus the theory offers an explanation of the results of numerous experiments showing that younger children fail to take such factors into account. (Gordon, 2009)

So a simulationist account has a nifty explanation of just which part of the process breaks down under cognitive loads, predicting roughly the data we see.

References

Behne, T., Carpenter, M., and Tomasello, M. (2005). One-year-olds comprehend the communicative intentions behind gestures in a hiding game. *Developmental Science*, 8(6):492–499.

Breheny, R. (2006). Communication and folk psychology. *Mind and Language*, 21(1):74–107.

Burge, T. (2015). Do apes and very young children attribute mental states? YouTube Video.

Buttelmann, D., Carpenter, M., and Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, 112:337–342.

Butterfill, S. A. and Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind and Language*, 28(5):606–637.

Carey, S. (2009). *The Origin of Concepts*. Oxford University Press.

Carruthers, P. (2013). Mindreading in infancy. *Mind and Language*, 28(2):141–172.

Csibra, G. (2010). Recognizing communicative intentions in infancy. *Mind and Language*, 25(2):141–168.

DeChant, R. (MS). Communication, mindreading, and cognitive architecture: Against sperber and wilson's case for a relevance-theoretic utterance interpretation module. *Unpublished Manuscript*.

Dennett, D. C. (1969). *Content and Consciousness*. Routledge.

Dennett, D. C. (1971). Intentional systems. *The Journal of Philosophy*, 68(4):87–106.

Dennett, D. C. (1987). *The Intentional Stance*. MIT Press.

Fodor, J. A. (1992). A theory of the child's theory of mind. *Cognition*, 44:283–296.

Gordon, R. M. (2009). Folk psychology as mental simulation. In Zalta, E. N., editor, *Stanford Encyclopedia of Philosophy (Fall 2009 Edition)*.

Happé, F. and Loth, E. (2002). 'theory of mind' and tracking speakers' intentions. *Mind and Language*, 17(1–2):24–36.

Joseph, R. M. and Tager-Flusberg, H. (2010). Preschool children's understanding of the desire and knowledge constraints on intended action. *British Journal of Developmental Psychology*, 17(2):221–243.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics*, 4(1):71–84.

Kikuno, H. and ans Fenja Ziegler, P. M. (2007). How do young children process beliefs about beliefs?: Evidence from response latency. *Mind and Language*, 22(3):297–316.

Leslie, A. M. (1987). Pretense and representation: the origins of 'theory of mind'. *Psychological Review*, 94:412–426.

Leslie, A. M. (1994). Tomm, toby and agency: Core architecture and domain specificity. In Hirchfeld, L. and Gelman, S., editors, *Mapping the Mind*, pages 119–148. Cambridge University Press.

Leslie, A. M., German, T. P., and Polizzi, P. (2005). Belief-desire reasoning as a process of selection. *Cognitive Psychology*, 50(1):45–85.

Luo, Y. (2011). Do 10-month-old infants understand others' false beliefs? *Cognition*, 121:289–298.

Moore, R. (MS). Gricean communication and cognitive development. *Unpublished Manuscript, Max Planck Institute for Evolu- tionary Anthropology.*

Onishi, K. and Baillargeon, R. (2005). Do 15-month-olds understand false beliefs. *Science*, 308:255–258.

Perner, J. and Ruffman, T. (2005). Infants' insight into the mind: How deep? *Science*, 308:214–216.

Roth, D. and Leslie, A. M. (1998). Solving belief problems: toward a task analysis. *Cognition*, 66:1–31.

Santana, C. (MS). Intention-based accounts of meaning are idealized scientific models. *Unpublished Manuscript.*

Sellars, W. (1956). Empiricism and the philosophy of mind. In Feigl, H. and Scriven, M., editors, *The Foundations of Science and the Concepts of Psychology and Psychoanalysis*, volume 1 of *Minnesota Studies in the Philosophy of Science*, pages 253–329. University of Minnesota Press, Minneapolis.

Southgate, V., Senju, A., and Csibra, G. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18(7):587–592.

Sperber, D. and Wilson, D. (2002). Pragmatics, modularity, and mindreading. *Mind and Language*, 17:3–23.

Thompson, J. R. (2008). Grades of meaning. *Synthese*, 161:283–308.

Thompson, J. R. (2014). Meaning and mindreading. *Mind and Language*, 29(2):167–200.

Wellman, H. M., Cross, D., and Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72(3):655–684.

Wimmer, H. and Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, 13:103–128.