Origins of Mind: Lecture Notes
Lecture 09

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

December 5, 2013

title-slide

Action

slide-3

Our first question is, When do human infants first track goal-directed actions and not just movements? In examining nonlinguistic communication, we've assumed that infants from around 11 months of age can produce and comprehend informative pointing. This commits us to saying that they have understood action.

slide-4

#source 'research/teleological stance – csibra and gergely.doc' #source 'lectures/mindreading and joint action - philosophical tools (ceu budapest 2012-autumn fall)/lecture05 actions intentions goals' #source 'lectures/mindreading and joint action - philosophical tools (ceu budapest 2012-autumn fall)/lecture06 goal ascription teleological motor' When do human infants first track goal-directed actions and not just movements? Here's a classic experiment from way back in 1995. The subjects were 12 month old infants. They were habituated to this sequence of events.

slide-6

There was also a control group who were habituated to a display like this one but with the central barrier moved to the right, so that the action of the ball is 'non-rational'.

slide-7

For the test condition, infants were divided into two groups. One saw a new action,
...

slide-9

... the other saw an old action. Now if infants were considering the movements only
and ignoring information about the goal, the 'new action' (movement in a straight
line) should be more interesting because it is most different. But if infants are taking
goal-related information into acction, the 'old action' might be unexpected and so
might generate greater dishabituation.

slide-13

You might say, it's bizarre to have used balls in this study, that can't show us anything
about infants' understanding of action. But adult humans naturally interpret the
movements of even very simple shapes in terms of goals. So using even very simple
stimuli doesn't undermine the interpretation of these results.

slide-14

Consider a related study by Woodward and colleagues. (It's good that there is con-
verging evidence from different labs, using quite different stimuli.)

slide-16

'Six-month-olds and 9-month-olds showed a stronger novelty response (i.e., looked
longer) on new-goal trials than on new-path trials (Woodward 1998). That is, like
toddlers, young infants selectively attended to and remembered the features of the
event that were relevant to the actor's goal.' (Woodward et al. 2001, p. 153)

slide-17

Consider a further experiment by Csibra et al. (2003). This is just like the first
ball-jumping experiment except that here infants see the action but not the circum-

stances in which it occurs. Do they expect there to be an object in the way behind that barrier?

slide-19

Why think about adults when you want to know about the origins of knowledge? Because sometimes it's possible to identify adult abilities which are plausibly identical to infant abilities. We saw an example of this in thinking about knowledge of minds, and in thinking about knowledge of colour. In both cases, linking infants' competence to adults' competence allowed us to better understand the competence. Can we also make such a link in the case of action?

slide-20

I already mentioned this classic study by Heider and Simmel. This kind of study has been done with human adults quite a bit since then. It is quite tempting to suppose that what we see here is automatic ascription of goals in adults.

slide-24

But is it automatic? Is it perceptual? Is there any evidence?

slide-25

Premack (1990) and Scholl and Tremoulet (2000) all draw a parallel between causation and action. (Premack's hypothesis that infants' understanding of intention is comparable to the understanding of causal interactions as shown by Michotte's launching stimuli ((Premack 1990); see also (Premack and Premack 1997)).) We saw (back in lecture 4) that there is quite convincing evidence that causal interactions are represented in perceptual proceses. So if there is really a parallel, we could infer that relations between actions and goals are represented in perceptual processes too.

slide-28

[Here I'm not trying to provide evidence, only to explain what the claim commits us to.] Recall this habituation experiment. they suggests that both

slide-29

The new Gergely et al experiment about action is based on the same habituation technique.

slide-33

If I had more time, I'd tell you a long story about speech and action. Maybe some of us will get together another time and I can tell you that story then. (*nb: (Zwickel et al. 2010) also argue for perceptual categories in perceiving actions, and do so by comparison with pop-out effects for orientations.)

slide-35

Subjects have to judge whether the dot is to the left or the right of the triangle (from their perspective). If you think this is just a triangle, then it doesn't have a left or right so there's no congruent or incongruent. But if you think the triangle performs goal directed actions, then in the figure on the top right, the dot is left of the triangle from your point of view but right of the triangle from its point of view. Could there be altercentric interference?

slide-38

Here are the results. When the triangle makes random movements: there's no difference in RT between congruent and incongruent conditions. (As you'd expect—this just shows that there's nothing wrong with the setup.) (The experiment involved comparing neurotypical (ordinary) subjects with AS subjects; that's interesting but too complex for us so we'll just focus on the neurotypical subjects' performance.) Could there be altercentric interference?

slide-41

What can we conclude? For adults, tracking information about goals makes one susceptible to interference from other other's perspective—it makes one susceptible to evaluting left and right from another's point of view. This altercentric interference effect is a long way short of showing that the goal ascription is automatic or perceptual. After all, such interference might be a consequence of thinking reflectively about the goals of the objects. But it is at least a step in the direction in that it shows

that, in adults, goal ascription can at least have effects on automatic processes. (If automatic processes are informationally encapsulated, then this is evidence that goal ascription is automatic.) (*Other evidence: (Gao et al. 2010; Teufel et al. 2010))

slide-42

Would be nice to talk about this but we don't have time (could use some Brian Hare experiments; also Gomez?.)

slide-43

(The three months of age figure comes from Sommerville and Woodward (2005).)m

slide-46

So it's maybe plausible there is core knowledge of action. What we haven't done here (and can't do, as far as I know) is to show that infants lack knowledge knowledge of action. So it's open to someone to deny that infants have core knowledge of action on the grounds that they have knowledge knowledge of it.

slide-49

Moving towards our next topic ... But, What do infants understand of actions and their goals? Or, as I like to put the quesiton, What model of action underpins infant cognition of action?

slide-50

What do infants understand of action? Intention? Premack (1990) says yes. But Geregely et al and Woodward et al say no.

slide-51

It isn't clear that this proposal can work (as introduced by Dennett, the intentional stance involves ascribing mental states), as these authors probably realised later, but the point about not representing mental states is good.

slide-53

What we've seen so far is (i) just what scientists think and (ii) almost entirely negative. It's that infants don't represent intentions but do represent actions and their goals. So we should ask, What do infants understand of actions and their goals? Or, as I like to put the quesiton, What model of action underpins infant cognition of action?

It's essential to see that there's a question here. We've shown that infants from around six months of age can track differences that are in fact differences in the actions. But we haven't shown how infants understand these differences; we haven't said how actions appear from the point of view of the infant. [***Could do something on which model here by analogy with the physical case.]

unit_711

Intentions and Goals

slide-55

In this section we ask, What do infants understand of actions and their goals? Or, as I like to put the quesiton, What model of action underpins infant cognition of action?

slide-56

What is the relation between an action and the outcome or outcomes to which it is directed?

slide-57

One feature of actions is that, among all their actual and possible outcomes, some are goals to which they are directed. I seize little Isabel by the wrists and swing her around, thereby making her laugh and breaking a vase. You might wonder what the goal of my action was. Did I act in order to break the vase or to make Isabel laugh? Or was my action perhaps directed to some other goal, one that was not realised so that my action failed? A basic challenge for an account of action is to explain the

relation between actions and the goal or goals to which they are directed. Among all of the actual and possible outcomes of an action, which are goals of the action?

slide-58

The standard answer to this question involves intention. An intention (1) represents an outcome, (2) causes an event; and (3) causes an event whose occurrence would normally lead to the outcome's occurrence. What singles out an actual or possible outcome as one to which the component actions are collectively directed? It is the fact that this outcome is represented by the intention. So the intention is what links an action to the outcomes to which it is directed.

slide-59

Now we have to ask, What is an intention?

slide-60

First note that goals are not intentions. Goals are actual or possible outcomes. They are states of affairs. Intentions, by contrast, are or involve mental states that represent a goal. It would be a terrible mistake to confuse a goal with a goal-state. That would be like confusing a person with a photograph.

slide-61

So what is an intention?

slide-62

[*new route: for the purposes of understanding these agents, we want the simplest possible notion of intention; so we'll take intention as an action-causing belief-desire pair. The idea I want to consider is, surprisingly, that there are no such things as intentions.

> 'The expression 'the intention with which James went to church' has the outward form of a description, but in fact it ... cannot be taken to refer to an entity, state, disposition, or event. Its function in context is to generate new descriptions of actions in terms of their reasons; thus

'James went to church with the intention of pleasing his mother' yields a new, and fuller, description of the action described in 'James went to church'.' (Davidson 1963, p. 690)

What motivates this view? We already have beliefs and desires in our model of action explanation. Introducing intentions as additional mental states would make the model more complicated. So if we can do without intentions, we should do so in the interests of simplicity.

slide-63

But how can we do without intentions? Haven't we just seen that we need intentions in order to explain the relation between an action and the goal or goals to which it is directed?

slide-64

Here's how Davidson's view works. James desired to please his mother. James believed that going to church would please his mother. And this belief and desire caused his going to church. So the belief–desire pair can play the role of an intention. It (1) represents an outcome—in this case, the pleasing of James' mother—, (2) causes an event—James' going to church—; and (3) causes an event whose occurrence would normally lead to the outcome's occurrence.

slide-65

It appears, then, that we can explain the relation between an action and the goal or goals to which it is directed just in terms of belief and desire. We don't need to introduce intentions as further mental states. If we like we can say that an intention just is a suitable, action-causing belief-desire pair. This view of intention is parallel to a view about knowledge. Some suppose that knowledge is justified true belief. Or belief meeting some condition like being true and justified. On this sort of view, knowledge is not a mental state over and above belief. Rather, there are just beliefs and some of these beliefs have a special status. Similarly, Davidson's idea might be put by saying that intentions are just action-causing belief–desire pairs. (This is not to say that *any* belief–desire pair is an intention, of course.)

slide-66

Let me give you one more example. Let's assume that we know what beliefs and desires are. It's a familiar idea that they combine to cause actions like this:

slide-69

To say that he had the intention to avoid rabies is, on this view, just to say that his action was caused by a belief and a desire with these contents.

slide-70

Now this isn't the right way to model adults' understanding of intention, but it's an approximation that works in the cases we're concerned with. (Interesting question is why we can't stop with intention = action-causing belief-desire pair and what intentions are for, but this is a topic for the course about reasons and actions.])

slide-71

Tomasello and Call say that this *is* the model of action that underpins primate cognition

slide-72

And we saw earlier that Premack says the same about infants, although not everyone agrees.

slide-73

So should we just stop here? There are at least two reasons not to stop but rather to look for alternative models of action.

slide-75

(a) It would be good to see if there is a simpler way of understanding this relation, perhaps one that doesn't involve insight into mental states. (b) Minimal theory of mind ... presupposes an understanding of goal-directed action. (You might say, on

minimal theory of mind that instead of beliefs and desires we could combine registration and preference; but that would affect the construction since registration was explained in terms of goal-directed action). So, is there a simpler model of the relation, one that infants might understand?

unit_712

Pure Goal Ascription: the Teleological Stance

slide-77

Goal ascription is the process of identifying outcomes to which purposive actions are directed as outcomes to which those actions are directed.

slide-79

Pure goal ascription is goal ascription which occurs independently of any knowledge of mental states. Here we consider two accounts of pure goal ascription, one involving the teleological stance the other involving motor cognition.

slide-81

In looking for an alternative model of action we are also asking, How could pure goal ascription work? Could you maybe think about the goal and not the intention and that would be enough? Well no because a goal is just an outcome. Let me try to explain with an example ...

slide-82

Earlier I said that *Goal ascription* is the process of identifying outcomes to which purposive actions are directed as outcomes to which those actions are directed. Given this definition, goal ascription involves representing three things:

1. an action

2. an outcome

and

3. the directedness of the action to the outcome.

It is important to see that the third item—representing the directedness—is necessary. This is quite simple but very important, so let me slowly explain why goal ascription requires representing the directedness of an action to an outcome. Imagine two people, Ayesha and Beatrice, who each intend to break an egg. Acting on her intention, Ayesha breaks her egg. But Beatrice accidentally drops her egg while carrying it to the kitchen.

slide-83

So Ayesha and Beatrice perform visually similar actions which result in the same type of outcome, the breaking of an egg; but Beatrice's action is not directed to the outcome of her action whereas Ayesha's is.

slide-84

Goal ascription requires the ability to distinguish between Ayesha's action and Beatrice's action. This requires representing not only actions and outcomes but also the directedness of actions to outcomes.

This is why I say that goal ascription requires representing the directedness of an action to an outcome, and not just representing the action and the outcome.

slide-85

So how could pure goal ascription work? How could we represent the directedness of an action to an outcome without representing an intention?

slide-86

Csibra and Gergely offer a 'principle of rationality' according to which ... Csibra & Gergely's principle of rational action: 'an action can be explained by a goal state if, and only if, it is seen as the most justifiable action towards that goal state that is available within the constraints of reality.'(Csibra and Gergely 1998; Csibra et al. 2003) (Contrast a principle of efficiency: 'goal attribution requires that agents expend the least possible amount of energy within their motor constraints to achieve a certain end.' (Southgate et al. 2008))

This principle plays two distinct roles. One role is mechanistic: this principle forms part of an account of how infants (and others) actually ascribe goals. Another role is normative: this principle also identifies grounds on which it would be rational to ascribe a goal.

As Csibra and Gergely formulate it, the principle might seem simple. But actually their eloquence is hiding some complexity.

slide-87

How are we to understand 'justifiable action towards that goal state?'

slide-88

It is perhaps worth spelling out what might be involved in applying this principle. Let me try to spell it out as an inference with premises and a conclusion ...

slide-92

An action of type a' is a better means of realising outcome G in a given situation than an action of type a if, for instance, actions of type a' normally involve less effort than actions of type a in situations with the salient features of this situation and everything else is equal; or if, for example, actions of type a' are normally more likely to realise outcome G than actions of type a in situations with the salient features of this situation and everything else is equal.

A problem with what we have so far is side-effects, which can be highly reliable. Actions typically have side-effects which are not goals. For example, suppose that I walk over here with the goal of being next to you. This action has lots of side-effects:

- I will be at this location.

- I will expend some energy.

- I will be this much further away from the front

These are not goals to which my action is directed. But they are things which my action would be a rational and efficient way of bring about. So there is a risk that these optimising versions of R will over-generate goals.

slide-93

I think this first problem can be solved by adding a clause about desire. [*] We can substantially mitigate the problem of side-effects by requiring that $R(a, G)$ hold only where G is the type of outcome which is typically desirable for agents like a.

slide-95

Now so far we've been considering this as an account of how someone could identify to which goal an action is directed without thinking about mental states. That is, this inference is the core component in an account of pure goal ascription.

slide-96

But insofar as we think of this inference as an heuristic for identifying goals, we haven't answered the question I started with. Using the inference presupposes that you already understand what it is for an action to be directed to a goal. But we can also turn it around and use it to define a relation between an action and a goal.

slide-97

So let's define a relation between an action and a goal, R(a,G), which obtains just when conditions (1)–(5) are met.

slide-98

(Do you remember Dennett's ingeneous twist in The Intentional Stance? We're making the same move here.)

slide-102

This is what we've just shown in arguing about (1)-(5)

slide-103

This seems to follow from the way it's defined.

slide-104

This is an issue I'll come back to later.

unit_741

From Action to Communication and Joint Action

slide-111

Note that Csibra & Gergely offer a continuity hypothesis.

slide-117

recall from last lecture ... Now contrast Grice and Davidson on the pointing action from the Hare et al study, where you're supposed to take one of two containers. Strictly speaking, that Ben should come over might not be the first meaning of the wave (so there are other options here).

slide-120

As before, there's a contrast in what must be intended and so what we're committing ourselves to in saying that infants can produce and comprehend informative pointing.

unit_999

Conclusion and Outlook

slide-126

How do humans come to know about—and to knowingly manipulate—objects, causes, words, numbers, colours, actions and minds? Have we made any progress?

slide-127

If you're wondering what the big idea is, there isn't one. I think progress depends on approaching it case by case. For instance, we considered the role of language in two cases, knowledge of colour and knowledge of minds. It seems to play quite different roles in these two cases. So while it seems plausible that abilities to communicate by language do play a role in explaining the emergence of knowledge in many domains, it's hard to assign it one role. There are themes running across the cases but little in the way of general principles. I'm not saying that's a bad thing; that's just how it is.

slide-128

Core knowledge is initially a label for a problem. With respect to various domains of knowledge including colour, physics and psychology, infants have some but not all of the capacities that are characteristic of knowledge. So we face a dilemma. The simplest explanation of what they can do would be to ascribe them knowledge. But ascribing them knowledge systematically generates false predictions about what they will do. So it must be wrong to ascribe them knowledge. In the first instance, □ core knowledge□ (or modular cognition, or implicit knowledge, or tacit knowledge) is just a label for whatever it is that is not knowledge but explains these capacities.

slide-129

Developmental evidence for core knowledge is fragile: it is always possble that the apparently incorrect predictions can be explained just by appeal to performance factors. Now as I understand the science many attempts to do this have failed. But for all we know this may be due only to the limited ingenuity of the researchers. Reflection on adults provides us with a second line of converging evidence. To say that a process is automatic is ... We're looking for two features in adults: (i) apparent conflict in what the adults know depending on whether the response is automatic or controlled, and (ii) limits of adults' automatic competence match those of infants'. We found some evidence this pattern in (a) colour; (b) mindreading; (c) number; (d) maybe causation and physical objects (representational momentum (Kozhevnikov and Hegarty 2001)).

slide-130

I'm illustrating this pattern with Low & Watts but be careful — first look is spontaneous but may not reflect automatic processing.

slide-133

This picture is significantly different from some competitors (but not Carey on number): (1) because it shows we aren't done when we've explained the acquisition of core knowledge (contra e.g. Leslie, Baillargeon), and (2) because it shows we can't hope to explain the acquisition of knowledge if we ignore core knowledge (contra e.g. Tomasello)

slide-135

Patterns in the mindreading case seem different from patterns in the cases of physics and number. Most people are committed to there being different stories for (a) colour (categorical perception); (b) syntax (tacit knowledge); and (c) physical objects. I'm not sure. I'm tempted to think that perceptual and motor cognition are at the root of these in every case, but this is a radical idea for which there is currently very little evidence.

slide-136

Conclusive evidence is rare, but that's not really an obstacle. There's no reason to take radical anti-nativism as our starting point. This question has two aspects: (i) what sort of thing; (ii) what particulars in which domains. On (i), I like the idea that core knowledge is innate. But we should be more careful and say at most that either core knowledge or something capable of priming the acquisition of core knowledge is innate. On (ii), we looked at a paradigm case, knowledge of syntax. In this case, we used a poverty of stimulus argument. While the evidence is less clear in other cases, there is certainly reason to suppose that poverty of stimulus arguments might work for knowledge of colour and knowledge of physical objects. But I also tried to argue that deciding this question probably isn't very important from our point of view.

slide-137

We've seen that cooperative forms of social interaction appear early in development. We studied pointing but there are examples involving helping and playing together too. In seminars we've been considering two ideas: (i) there is a form of cooperative social interaction abilities to engage in which are important for explaining the origins of knowledge of others' perspectives, for explaining the emergence of abilities to communicate by language, and norms. (ii) these forms of cooperative social interaction involve a capacity for shared intentionality, where shared intentionality is understood as the ability to share mental states.

slide-138

We should characterise the abilities that underpin cooperative social interactions in infancy without appeal to romantic but mysterious ideas about sharing.

slide-139

Need to re-think what's needed for social interaction if it's to explain things, And the starting point for this is to reflect on abilities to understand action.

slide-140

I like this picture: core knowledge + social interaction –> abilities to communicate –> knowledge knowledge Two problems: (i) lack of evidence; (ii) creativity in acquisition of communication by language is hard to square with the prior absence of knowledge knowledge.

slide-141

I've done one of these pictures for colour and for knowledge of syntax. This is speculative.

slide-142

slide-143

Well, maybe there is a big idea after all. (I'm hesitant to mention this because it's not well established and, anyway, you should find your own big idea. But ...) Through social interaction you are able to rediscover what is in some sense already encoded in your core knowledge. That, anyway, is one idea about how humans come to know about objects, causes, minds and the rest.

References

Csibra, G., Bíró, S., Koós, O., and Gergely, G. (2003). One-year-old infants use teleological representations of actions productively. *Cognitive Science*, 27(1):111–133.

Csibra, G. and Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science*, 1(2):255–259.

Davidson, D. (1963). Actions, reasons, and causes. *The Journal of Philosophy*, 60(23):685–700.

Gao, T., McCarthy, G., and Scholl, B. J. (2010). The wolfpack effect perception of animacy irresistibly influences interactive behavior. *Psychological Science*, 21(12):1845–1853. PMID: 21078895.

Kozhevnikov, M. and Hegarty, M. (2001). Impetus beliefs as default heuristics: Dissociation between explicit and implicit knowledge about motion. *Psychonomic Bulletin & Review*, 8(3):439–453.

Premack, D. (1990). The infant's theory of self-propelled objects. *Cognition*, 36(1):1–16.

Premack, D. and Premack, A. J. (1997). Infants attribute value[+-] to the goal-directed actions of self-propelled objects. *Journal of Cognitive Neuroscience*, 9(6):848–856.

Scholl, B. J. and Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, 4(8):299–309.

Sommerville, J. A. and Woodward, A. L. (2005). Pulling out the intentional structure of action: the relation between action processing and action production in infancy. *Cognition*, 95(1):1–30.

Southgate, V., Johnson, M. H., and Csibra, G. (2008). Infants attribute goals even to biomechanically impossible actions. *Cognition*, 107(3):1059–1069.

Teufel, C., Fletcher, P., and Davis, G. (2010). Seeing other minds: attributed mental states influence perception. *Trends in cognitive sciences*, 14(8):376–382.

Woodward, A. L., Sommerville, J. A., and Guajardo, J. J. (2001). Making sense of human behavior: Action parsing and intentional inference. In Malle, B. F., Moses, L., and Baldwin, D. A., editors, *Intentions and Intentionality*, pages 149–169. MIT Press, Cambridge, MA.

Zwickel, J., Grosjean, M., and Prinz, W. (2010). On interference effects in concurrent perception and action. *Psychological Research*, 74(2):152–171.