A concise phenomenology of folk psychology

June 29, 2020

1 Introduction

Begin with Heyes' view on the situation in animal mindreading research: there's a methodological problem and a theoretical problem. The second is the most urgent: "there is difficulty in conceptualising alternatives to 'full-blown' mindreading", and as a consequence "it is no longer clear what the search for animal mindreading is searching for."

But what is "full blown mindreading" (with or without scare quotes)? Heyes doesn't say, and although she observes that "there is no proprietary definition of what it is to engage in mindreading or to have a theory of mind", she doesn't seem to be worried by this. But if the problem lies in conceptualising alternatives to full-blown mindreading, shouldn't we first clarify what full-blown mindreading is? (Ditto for infant studies.)

Likewise for minimal theory of mind, which according to Heyes "provides an exceptionally promising potential solution to the theoretical problem at the heart of contemporary research on animal mindreading." How can we say that MTM is a theory of mind at all unless it has been established what a full-blown theory of mind is? (Ditto for other subnormal models of mind.)

So that's the question we will be addressing here. Narrow this down to belief attribution.

The approach. The bulk of the literature on mindreading, ToM, etc. treats these as labels of a psychological capacity (or, perhaps, an ensemble of psychological capacities). Although ToM is now often used as a descriptive term ("social cognition"), it originally implied that this capacity involves a theory; other accounts deny this. Our approach is descriptive rather than theoretical. Our goal is to describe what people actually do when they attribute

mental states to others or themselves, rather then develop theoretical hypotheses about the psychological capacities that enable them to attribute mental states. Briefly: we will be concerned with certain behaviours that people engage in. Call this "folk psychology". FP is a social practice, which manifests a psychological capacity, which we will refer to as "mindreading" (or maybe "social cognition"?). Full-blown mindreading must at least be able to support this social practice.

The social practice we have in mind is overt mental-state attribution. If Fred says, "Betty thinks that Barney snores", then it is practically certain that Fred attributes to Barney the belief that Betty snores, and the evidence won't get more direct than that. Overt mental-state ascription is the paradigm, and covert (or "implicit") mental-state ascription is modeled after it. We will focus on belief ascription, but much of what we will have to say about that generalises to other propositional attitudes.

Clarify the relation between folk psychology and mindreading by means of addition (inspired by Marr): addition as an arithmetical operation vs. cognitive processes for adding numbers. Or traffic vs. the psychology of road users.

Once we have a description of overt belief attribution, we can turn to chimpanzees and human infants, and consider *in what respects* and *to what extent* the capacities they exhibit in experimental settings meet the requirements for overt belief attribution. Similarly, we will be in a position to provide a detailed characterisation of sub-normal models of mind in relation to full-blown mindreading.

Isn't it naive to think that a purely descriptive account of overt mental-state attribution is possible at all? We don't think so. Of course, there is no guarantee that our account is correct, but still our enterprise is a lot less conjectural than theory formation. First, our database is vast. People produce overt mental-state attributions on a regular basis, and a signficant portion of them are recorded in linguistic corpora, of which the World Wide Web is far and away the largest. Secondly, all of us are experts in language interpretation, which is the chief skill that we will be relying on. Thirdly, we will be careful not to make any controversial claims, and there will be nothing in our account that hasn't been said before, and is widely agreed to be correct.

2 Preliminaries

1. Since all or nearly all words are polysemous, it is practically an *a priori* truth that mental-state words are polysemous, too. (Cf. Borg et al. on "pain".)

   A nice paradigm is the word "school", which according to the OED has the following senses (among others):

   (a) an institution for educating children
   (b) the buildings used by a school
   (c) the pupils and staff of a school

   If I say, "The headmaster addressed the whole school", then (c) is the relevant sense, but note that this doesn't mean that (a) and (b) are simply discarded. At least (a) is part of the message, too. The various senses of a polysemous word are *connected*, and if you select one, others may come along.

2. In addition to polesemy, there is prototypicality. Content-words (or the concepts associated with them) are associated with "prototype effects", which may depend on the context. E.g., prototypical red is different for wine, hair, skin,...

3 Phenomenology and diagnostics

P_1 Many though not all of the mental states that we attribute to one another appear to have a species of content that agrees with reality or not. Beliefs are true or false; if I intend to do the dishes, then I may or may not realise my intention by acting accordingly; and so on.

Mental states that don't have content in this sense (or not necessarily) are nervousness, anger,...

Contents may be more or less complex along several dimensions.

(a) The content of a mental state may involve another mental state. Higher-order mental states. Higher order is more complex, but not necessarily harder. Cf. complexity vs. difficulty of number words within and beyond the subitising range.

(b) *De re, de dicto*, or plain (both). Pre-theoretically, plain attribution seems to be the most basic and simplest. (Loar 1972, Geurts 1998) In a sense, this is a conflation of *de re* and *de dicto*, but that doesn't mean it is more complex than either. Compare Dutch "neef", which applies to nephews and cousins alike, or languages that use the same word for blue and green.

Conflation of *de re* and *de dicto* may have the same source as (or be related to) the conflation of content and form. "Pigs must be called pigs because they *are* pigs."

(c) In some cases, (the content of) a mental state has a "direction of fit". Belief and intention are among the clearer cases.

Diagnostics for belief attribution:[1]

∗ If A believes that S is true/false, then A considers B's belief to be true/false.

∗∗ A attributes to B the belief that C is rich, even though A doesn't believe that C exists. (*de dicto*)

A attributes to B the belief that C believes/wants/intends. . . (higher-order mental states)

A attributes to B the belief C is a fool but not the belief that D is a fool, even though A believes that C = D. (*de re*)

P_2 [Behavourial effects] Mental states guide our actions, and there are more or less systematic connections between mental states, on the one hand, and patterns of behaviour, on the other.

Corollary: things with mental states are things that act.

Diagnostics for belief attribution:

∗ A attributes beliefs to creatures that act (as opposed to, e.g., shoes, toothpaste, corpses).

A attributes to B the belief that the tea water is boiling / that the mushrooms on her plate are poisonous / that her husband is dead . . .

---

[1]Let's say that A attributes to B the belief that S iff A is disposed to affirm that B believes that S, where S is a declarative sentence that represents the content of the attributed belief. Hence, it is presupposed that belief attribution *must* involve truth-valued content (P_1). Knowledge is a form of belief.

P$_3$ Mental-state attribution (or the mental states themselves?) obeys certain rules/regularities: seeing is believing, inertia of belief, elementary logic, interactions between mental states (e.g. belief and intention),...

Part of this systematicity involves content.

There seem to be two forms of belief inertia. If you believe at noon that it is raining, then ceteris paribus you will believe tonight that it was raining *at noon*; but you won't necessarily believe tonight, even ceteris paribus, that it is *still* raining. (Compare this with your belief that Angela Merkel is German.) So, once you have formed a belief, you will stick with it, ceteris paribus; that's belief inertia proper. The other is that, based on world knowledge, you suppose that certain states of the world are inert, ceteris paribus, and your beliefs reflect these inertia assumptions.

Diagnostics for belief attribution:

* Seeing is believing, inertia$_1$, inertia$_2$, modus ponens,...
** Harder forms of reasoning

P$_4$ Mental states are private.

P$_5$ Mental-state attribution is normative. (Always?)

P$_6$ Mental states are somewhere. In our culture they are generally between the ears. Exceptions are pains, itches,...

4 Discussion/applications

1. Mental states of infants, pets, insects, trees, artefacts, robots.
2. Standard false-belief task. (Importance of inertia.)
3. "Implicit" false-belief tasks.

5 Thinking things

1. "my husband thinks": 610.000 hits
   "My husband thinks he is always right."

2. "my dog thinks": 363.000 hits
   "My dog thinks I'm a genius."

3. "my phone thinks": 121,000 hits (cf. "my phone says": 777.000 hits)
   "My phone thinks I'm in Canada."

4. "my cat thinks": 108,000 hits
   "My cat thinks he's human."

5. "the bees think": 57,100 hits
   "The bees think it's June."

6. "the thermostat thinks": 4,600 hits
   "With the module set to the middle setting, the actual room temperature would be about 62 degrees while the thermostat thinks it is 72 degrees."

7. "my calendar thinks": 300 hits (cf. "my calendar says": 147.000 hits)
   "My calendar thinks it is May 4th, how do I reset it to May 9th?"

8. "my pencil thinks": 50 hits
   "Too bad my pencil thinks that it's a helicopter and keeps flying away."

	...knows	...thinks	...believes	...wants	...intends
he...	127905	78569	75208	198948	9853
she...	55684	32600	28997	92891	2841
it...	14109	8378	9007	37870	6216

https://www.english-corpora.org/iweb/

1. the object named "player", which has been created by *the game* and contains everything *it knows* about the player character

2. *the team* is sure that *it knows* how to catch up

3. Now when *the control software* reads G28 in a line then *it knows* to send the machine to this position

4. Yeah, or someone would put a weighted dummy sitting in *the seat* so *it thinks* there is always a passenger

5. as *the flight controller* fights to hold the craft in what *it thinks* is the correct attitude

6. *the company* said *it thinks* Whole Foods shares are undervalued