A concise phenomenology of folk psychology

May 22, 2020

1   Introduction

There has been a lot to do about folk psychology, ToM, etc., and more in particular about (false) belief attribution. This is supposed to be something that people do all the time, and there is no shortage of theories about this sort of activity. But strangely enough, there have been no serious attempts at *describing* FP in any detail. Researchers have been theorising and arguing about experimental data, using words like "belief", "intention", etc. without considering in earnest what these words mean.

This might not seem to be problematic. After all, we generally take the meanings of our words for granted, and rarely stop to clarify the meaning of "milk" before announcing that we've run of milk. However, things are different if the *applicability* of a word is at issue (Is coconut milk really milk?), as it often is in discussions about FP: Do 15-month olds or chimpanzees really attribute beliefs to each other? Does level-1 mind reading involve true belief attribution or not? Etc.

The general strategy in discussions like these is to take the meaning of "belief" for granted, and rely on a small number of diagnostic criteria — often only one: the false-belief task. This is problematic not only because it is unprincipled (why the false-belief task?), but also because it is irresponsible to take it for granted that there must be a dichotomy between positive and negative instances of "belief". To see why, consider our everyday practice of attributing colours to things: "This is red", "That is green", and so on. This, too, might seem unproblematic until we come to realise that the meaning of a colour word varies with the context of use: "red" has clearly different ranges of applicability depending on whether we are discussing tomatoes, wine, skin, hair, grapefruits, or political doctrines. This context dependence is a *normal* feature of words; it is hard to find any words that don't have it. Therefore it is a moral certainty that the meaning of "belief" depends on the context, as indeed the OED confirms it does (8 basic uses).

This appears to be a practical problem in understanding research on ToM:

> 'chimpanzees understand ... intentions intentions ... perception perception and knowledge ... Moreover, they understand how these psychological states work together to produce intentional action 'chimpanzees probably do not understand others in terms of a fully human-like belief–desire psychology' (Call & Tomasello, 2008, p. 191).

We are not the first people to have noticed that there is a problem:

> 'the core theoretical problem in contemporary research on animal mindreading is that ... the conception of mindreading that dominates the field ... is too underspecified to allow effective communication among researchers, and reliable identification of evolutionary precursors of human mindreading through observation and experiment' (Heyes, 2015, p. 321)

One way to avoid the problem altogether is to frame debate in terms of tracking. A process *tracks* beliefs if how that process unfolds nonaccidentally depends, perhaps within limits, on facts about beliefs. Where there is tracking, you can change how the process unfolds by changing facts about what is believed. One advantage of focussing on tracking is that it allows us to stay closer to the observations than does talk about belief attribution. But to focus exclusively on tracking would be unsatisfactory. After all, the question was supposed to be whether chimpanzees (or infants, or ...) have a theory of mind, where to have a theory of mind practice is to be able to attribute mental states. And to answer any such question, we will need a clearer idea of what this ability is than can be provided by uncritical reliance on everyday mental state terms like 'belief' and 'intention'.

Our aim is to describe our everyday practice of attributing mental states (notably beliefs and intentions) in a relatively theory-neutral way. We believe that this project is a feasible one (if it is doable for redness, why not for belief?), and that it is theoretically important, because it will help to put debates about FP on a more solid footing.

The false belief task (**?**) is sometimes regarded as an acid test of mental state representations (see Bennett's, Dennett's and Harman's influential responses to **?**). But as **?**, p. 622 suggested, a false belief task is 'another arrow worth having in one's quiver rather than the assured bullseye that the philosophers suggest it is.'

2   Preliminaries

1. Separate social practices from capacities and processes.

2. Our primary target is the explicit attribution of beliefs and intentions. This is a vernacular practice, and this is where words like "believe" and "intend" get their original meaning.

   Arguably, *implicit* attribution of mental states is part of this practice, too. E.g., if Steve says, "Tomorrow is Tuesday", then normally speaking Bart is thereby licensed to assume that Steve believes that tomorrow is Tuesday. (This much should be acceptable regardless what your views on the pragmatics of assertion are.)

3. Theories about capacities and/or processing often use mental-state expressions from the vernacular practice without much discussion of their semantics. This raises the question what such expressions might mean, in the research of a given research project.

4. 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.

5. 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*)

P₂ [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.

---

[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 (P₁). Knowledge is a form of belief.

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 . . .

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.

References

Call, J. & Tomasello, M. (2008). Does the chimpanzee have a theory of mind? 30 years later. *Trends in Cognitive Sciences, 12*(5), 187–192.

Heyes, C. M. (2015). Animal mindreading: what's the problem? *Psychonomic Bulletin & Review, 22*(2), 313–327.