

The distinctive character of knowledge

ABSTRACT: Because knowledge entails true belief, it is can be hard to explain why a given action is naturally seen as driven by one of these states as opposed to the other. A simpler and more radical characterization of knowledge helps to solve this problem while also shedding some light on what is special about social learning.

Knowing that something is the case is not the same as merely being right about it. The target article offers substantial evidence that knowing is easier to recognize than the state of just having a belief aligned with reality, but we need a sharper picture of knowledge to explain why this is so.

The authors characterize knowledge through a list of four features, starting with factivity, identified as the condition that “you can only know things that are true.” Understood this way, factivity does not distinguish knowledge from true belief, although the next feature stipulates that knowledge is not just true belief. Knowledge has something extra which is lacking in beliefs that are “true by coincidence,” but what? The two other listed features of knowledge—that others can know things you don’t, and that it is not modality-specific—are unhelpful. Others can have true beliefs you don’t, and belief is equally amodal. These four features are supposed to distinguish knowledge from other states, and the article promises to focus on “instances of mental state representations that have these four signature features.” Given that knowledge attribution is supposed to be simpler than belief attribution, mindreaders will presumably not invoke all four features explicitly, conceptualizing the witnessed state as “not just true belief,” and so on. But it’s unclear what part these features play in any given attribution.

The puzzle deepens when we focus on an ambiguity in mindreading tasks. If some desired object is in a drawer, we expect similar reaching behaviour from the agent who knows that it is there

and the agent who just has a true belief. The experimental subject's correct anticipation of that reach does not on its own reveal which state was attributed, if any. Some key studies that the article cites as supportive of easy knowledge attribution actually label that condition as true belief (e.g. Krachun, Carpenter, Call, & Tomasello, 2009). Casually, theorists often gloss the true belief label in terms of knowledge, for example describing the "true belief" condition as consisting in "cases where the [observed agent] knows the peg has moved" (O'Connell & Dunbar, 2003, 134). If there is a crucial difference between knowledge and true belief attribution, what explains these patterns of labelling and explaining, and which type of attribution is actually happening?

Theorists are not strictly wrong to label a knowledge condition as true belief, assuming the standard philosophical view that knowledge entails true belief. When someone knows that the peg has moved, it is true both that the peg has moved, and that the agent believes the peg has moved. The target article defends a non-standard view according to which it is possible to know without believing (§5.2), on the basis of intuitive responses to cases in which someone can barely remember a fact. However, the standard view could be retained and the article's overall thesis better supported by a performance error explanation of these borderline cases: they are positively classified by easier processes of knowledge attribution but not by harder processes of belief attribution. Now, even if theorists are correct that their knowledge conditions are true belief conditions, this is not to say that experimental subjects are equally indifferent. I agree that it is actually knowledge rather than true belief that is ordinarily attributed in simple control cases, but to defend this position we need to explain the difference in mentalizing.

Here's one proposal: knowledge is simply a factive mental state, where the factivity condition is read as *necessarily* binding agents only to truths, while true belief combines a pair of conditions, one of which (truth) is not mental, and the other of which (belief) is a liberalization of knowledge (Williamson, 2000). We originally track knowledge because the problem of learning what other agents have in mind comes bound with the problem of learning about the larger world; knowledge attribution then works as a special part of the solution to that larger problem. Because other agents have different viewing angles and histories of experience with objects in the shared environment, recognizing signs of their knowledge constitutes a distinctively powerful way of learning about reality. Watching someone who knows where the peg has moved can tell you where the peg is now, assuming we can identify them as knowing, for example through gaze cues. Rather than just being "useful for determining who can accurately inform you about where to look," or as input to calculations about agents' future reliability, factive mentalizing provides a more direct way of learning how things are in the world. Gettiered agents will also inform you accurately, but unlike knowing agents, they are not accurate in virtue of the basic type of mental state they have (belief), so recognizing their mental state does not license a direct updating of one's model of reality. When a cage contains a hidden zebra and a deceptively painted donkey, the agent who sees only the donkey will tell you that there is a zebra in the cage, but that agent's mental relationship to reality is not of a type that necessarily reflects the truth.

Detecting true belief as such, meanwhile, requires separate steps of mentalizing and figuring out what is happening in the world, because belief is defined by relaxing the factivity condition on knowledge. As theoreticians we are free to execute these steps separately, but we should not assume that experimental participants are doing so. Because belief is a liberalization of knowledge, a wider

array of conditions can produce it, so the relevant patterns are harder to learn. Belief-detection patterns are largely derivative of knowledge-detection patterns: for example, in the unwitnessed transfer task, the false belief is that an object is in a location where it was recently known to be (Nagel, 2017). This derivative character of belief attribution explains why some mindreaders attribute knowledge but not belief, while none seem capable of attributing belief but not knowledge. The derivative character of belief itself explains why theorists naturally and appropriately explain "true belief" control conditions in terms of knowledge.

Acknowledgements:

Thanks to Evan Westra and Timothy Williamson for comments. My research is funded by the Social Sciences and Humanities Research Council of Canada (Insight Grant #502484) and the Schwartz Reisman Institute of the University of Toronto.

References:

- Krachun, C., Carpenter, M., Call, J., & Tomasello, M. (2009). A competitive nonverbal false belief task for children and apes. *Developmental Science*, 12(4), 521-535.
- Nagel, J. (2017). Factive and nonfactive mental state attribution. *Mind & Language*, 32(5), 525-544.
- O'Connell, S., & Dunbar, R. (2003). A test for comprehension of false belief in chimpanzees. *Evolution and Cognition*, 9(2), 131-140.
- Williamson, T. (2000). *Knowledge and its Limits*. New York: Oxford University Press.