Apes track false beliefs but might not understand them

Outlook article for *Learning and Behavior* (2017)

Kristin Andrews (York University)

Summary: Apes can correctly determine how to help a person with a false belief. But they may not need a concept of *belief* to do so.

The recent study by Buttelmann and colleagues (2017) is the first interactive test of whether apes can respond appropriately when a social partner has a false belief. While this work, like gaze tracking studies (Krupenye et al. 2016), finds that apes can anticipate people’s actions and goals when they have a false belief, we still don’t know *how* the apes do so. Buttelmann et al. claim that they offer converging evidence supporting Krupenye et al.’s claim that apes understand other agent's beliefs While these studies are consistent with the claim that great apes understand beliefs, other mentalistic explanations are consistent with both findings.

The 2017 study is a modification of the active helping task initially developed for 18-month-old infants (Buttelmann et al. 2009). The current experimental situation goes like this. The apes were exposed to two boxes, and with the help of a human assistant, learned that the boxes can be locked by sliding a bolt on the front. After demonstrating competence unlocking the boxes, an experimenter entered the scene carrying an object, and showed the object to the ape. The experimenter opened one box and placed the object inside and left the room.

In the true belief condition, the experimenter returned to the room to witness the assistant move the object from its original location to the other box. Importantly, the experimenter looked away when the boxes were locked and unlocked, so the ape could infer that the experimenter didn’t know how to unlock the box.

In the false belief condition the experimenter remained outside the room while the assistant “sneakily” moved the object to the other box, and locked it.

Next, in both conditions, the experimenter tried to open the now-empty locked box. He was unable to open the box (since he didn’t know how to unlock it), made a helpless gesture, and then pushed the tray of boxes toward the ape. As in the test with the infants, the apes opened the box containing the object significantly more often in the false belief condition than in the true belief condition. Both populations appeared to be able to determine the experimenter’s goal in the false belief condition, namely that the experimenter wanted the object. In the true belief condition, they found a difference between the children and the apes. Children opened the empty box significantly more often in this condition, which is consistent with them realizing that the experimenter had the goal of opening the empty box. Apes, however, were at chance in the true belief condition. Nonetheless, they behaved differently in the two conditions, suggesting that a familiar goal—gaining an object—was easy to track even when the social partner had a false belief. (Having the goal of opening an empty box might be less salient to the apes.) In a second experiment researchers again found a distinction between apes’ performance in an ignorance and a false belief condition.

These results are offered as evidence that the apes know the content of the experimenter’s belief about the location of the object in each condition. Using that belief attribution, and seeing the experimenter’s action, apes and children are able to work out the missing part of the causal story, namely what the experimenter desires. The kind of content being attributed to the ape who passes this false belief task is something like this: *The experimenter is trying to open the yellow box. He believes that the object is in the yellow box, but his belief is false; the object is really in the blue box. Why would he try to open the yellow box? Well, since people who open boxes that they believe contain an object often do so because they desire the object, he must want the object. I’ll get it for him.*

This line of reasoning can explain the apes’ behavior in the false belief condition, but if something like that kind of content is what is being attributed to the apes, we need to know a bit more about the concepts being ascribed. What are we claiming when we say apes understand beliefs?

Though we’ve been asking for 40 years whether apes can attribute beliefs, we haven’t said much about what beliefs are other than that they are truth functional representations—they can be true or false. In contemporary philosophy, there are various accounts of belief, but there are three properties of beliefs that most of these accounts agree on.

First, a belief is understood as a type of propositional attitude. A proposition, or the meaning of a sentence, can be the object of a variety of attitudes—a belief, which takes the content to be true; a desire, which takes the content to be wanted; a fear, which takes the content to be frightening; etc.. Thus, the content of a belief attribution can be given sentential structure, even if it is mentally represented via a nonsentential vehicle.

Second, a belief attribution has the logical property of referential opacity. It is a logical truth that if Loeberfeld is the tooth fairy, then if Loberfeld weighs 150 lbs, so does the tooth fairy. But, if you don’t know that Loberfeld is the tooth fairy, then knowing that Loberfeld weighs 150 lbs tells you nothing about the tooth fairy, and it would be an error in attribution to ascribe to you the belief that the tooth fairy weighs 150 lbs. A correct belief attribution captures *how* the target is representing the information, and for this reason a belief attribution has certain affinities with level-2 perspective taking (something we don’t currently have evidence for in great apes (Karg et al. 2016)).

Third, on the most common accounts, a belief, along with a desire, cause people to act. But to know how a belief plus a desire will cause someone to act, one needs additional information about how belief and desire sets interact to cause a particular behavior. This information may be gained through experience, or it may be generated by thinking about one’s self.

As a referentially opaque propositional attitude that can be true or false, and that, along with desire, causes behavior, a belief is going to be something that requires cognitive sophistication to use and to understand. Given the high cognitive demands, some critics of the false belief task will claim that the apes can solve these sorts of problems by sub-mentalizing or via a theory of behavior, rather than by mental state attribution. But any account that describes the apes as unable to understand the narrative structure of the situation, as not understanding others as agents with some mental states—emotions, goals, knowledge, perceptual states, etc.—should be rejected. Given what else we know about great apes’ abilities in these areas, some of which are reviewed in the Buttelmann paper, such an explanation is implausible.

There is a middle ground between belief attribution and sub-mentalizing. Apes may be using non-propositional mental state attribution in their social interactions.

In the active helping task, apes can work out the experimenter’s goal by understanding that people seek out things where they last saw them, so in the false belief condition, the ape thinks the researcher’s goal is to acquire the object because the researcher is trying to open the box where he last saw it. The content of the ape’s reasoning would go like this: *He is trying to open the box he left the object in. Since people look for things where they left them, he must want the object. I’ll get it for him.*

The alternative hypothesis is a mentalistic hypothesis, because it appeals to the intentions of the actor, but it is mentalistic hypothesis that doesn’t involve propositional attitudes, because the attributed content isn’t opaque and it doesn’t take sentential structure. It doesn’t matter how the experimenter represents the object and its location; it only matters that the ape knows the experimenter’s goal, something the authors think we already have evidence for given past ape studies. This hypothesis also can explain the infant findings in Buttelmann et al. 2009.

Recall the belief content ascribed to the ape who responds correctly. In order to work out the experimenter’s goal in the false belief condition, the ape has to first seek an explanation in order to uncover the corresponding desire that causes the behavior, and the attribution of desire is used to discover the experimenter’s goal. I have argued that belief attribution is more sophisticated a tool than we need for almost all related studies (Andrews 2012). Belief attribution, given its logical properties, is best suited to explaining discrepant behaviors, rather than predicting quotidian ones.

Andrews, K. (2012). *Do Apes Read Minds? Toward a New Folk Psychology*. Cambridge, MA: MIT Press.

Buttelmann, D., Buttelmann, F., Carpenter, M., Call, J., & Tomasello, M. (2017). Great apes distinguish true from false beliefs in an interactive helping task. *PLOS ONE*, *12*(4), e0173793.

Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*(2), 337–342.

Karg, K., Schmelz, M., Call, J., & Tomasello, M. (2016). Differing views: Can chimpanzees do Level 2 perspective-taking? *Animal Cognition*, *19*, 555–564.

Krupenye, C., Kano, F., Hirata, S., Call, J., & Tomasello, M. (2016). Great apes anticipate that other individuals will act according to false beliefs. *Science*, *354*(6308), 110–114.