Spotlight, *Trends in Cognitive Sciences*, 17 November 2016.

Apes submentalise

Cecilia Heyes

All Souls College and Department of Experimental Psychology

University of Oxford

Oxford OX1 4AL

United Kingdom

Email: cecilia.heyes@all-souls.ox.ac.uk

Abstract

Making subtle and extensive use of eye-tracking technology, Krupenye and colleagues showed that, like human infants, great apes – chimpanzees, bonobos and orangutans – can accurately anticipate the goal-directed behaviour of an agent that holds a false belief. How do they do it, by mentalising or by submentalising?

Keywords

Mentalising; submentalising; mind reading; theory of mind; eye-tracking; apes.

Humans often predict and explain each other's behaviour by ascribing mental states. At the movies, we expect the spy to head for the desk because he *wants* the documents and *believes* they are hidden there. For nearly 40 years, biologists and psychologists have been trying to find out whether this capacity for 'mentalising', 'mind reading' or 'theory of mind' is shared by other apes [1-3]. Krupenye and colleagues recently reported a breakthrough in this quest – evidence that chimpanzees, bonobos and orangutans can attribute false beliefs [4].

In their study, apes watched movies. Here is a description of what happened in one of the movies (*4*, Experiment 2, FB2), the kind of description provided by Krupenye and colleagues, and that most people would give if asked by a friend: A human actor, behind bars and dressed in green, had a contest with another actor, in the foreground and dressed in a King Kong (KK) suit. On three successive occasions the green actor tried to get possession of a brick, but each time KK snatched the brick and hid it in one of two boxes on his side of the bars. On the first two occasions, the green actor patiently retrieved the brick from the box where it had been hidden. On the third occasion, the green actor left the room after KK had hidden the brick. While he was away, KK transferred the brick to the other box, then removed it from the second box and left the scene taking the brick with him. When KK had gone, the green actor returned to the room, took up the central position from which he had begun to retrieve the brick on previous occasions, and the movie stopped.

Krupenye and colleagues found that, when the action stopped, the apes' first eye movements tended to be in the direction of the box where the brick had been hidden before the green actor left the room. When the action is described in a familiar, folk psychological way (above) it is natural to interpret this as a sign of mind reading - that the apes expected the green actor to search the box where he falsely believed the brick to be hidden – or, at

minimum, as a sign of 'behavior reading' – that the apes expected the green actor to search the location where he had last seen the brick [4, p. 113].

This study is important because it pioneers the subtle use of eye-tracking to test for mentalising in animals. An infrared eye-tracker was used, without head restraint, not merely to check that the apes looked at the stimulus display, but to record how much they looked at four critical areas in seven phases of the action (4, Figures S3 and S4). Consequently, we can be more confident than in any previous study that key elements of the action sequence 'went in' to the minds of the apes. For example, the eye-tracking data indicate that the apes watched the movements of the brick when the green actor was out of the room, and therefore make it unlikely that, when he returned, the apes merely looked towards the location where *they* believed the brick to be hidden.

So, the study by Krupenye and colleagues set a new methodological standard for research on mentalising in animals, and showed something truly interesting about the apes – that they "accurately anticipated the goal-directed behaviour of an agent who held a false belief" [4, p. 113] – but it did not tell us whether the anticipation was due mentalising or to 'submentalising'; prediction of behaviour by low-level, domain-general psychological processes [5].

Rather than viewing the movie as a narrative in which agents acted on objects for reasons, the apes may have selectively encoded relatively low-level properties of salient events, including the appearance and disappearance of the striking green shirt; the configuration of three cues (green centre / bell rings / boxes flash) that signalled an excitingly novel event (the box taking flight; see Figure 1B); and a predictor of which box would fly next – the last location of the brick when the scene was green. Once the possibility of low-level encoding is acknowledged – once we have entered this potential "self-world" of the apes [6] - it becomes apparent that a range of domain-general mechanisms, that process

inanimate as well as animate stimuli, could have driven the apes' eye-movements. For example, reappearance of the green shirt could have acted as a retrieval cue, activating a memory of the brick's location when the green colour was last present [7]. Eye-tracking can give a fair indication of what goes in to a mind, but it can't tell us what is likely to come out - to be remembered. Alternatively, the orientation of the green object relative to the boxes and the brick prior to the green object's disappearance could have acted as a contextual cue priming the apes' visual search when the green object reappeared. Experiments using inanimate stimuli with adult human participants have shown that this kind of cueing is driven by incidentally learned associations between spatial configurations and target locations [8].

The retrieval cue and contextual cue hypotheses suggest that the apes were submentalising – predicting behaviour using mechanisms of attention, learning and memory that did not evolve for, and are not dedicated to, the analysis of agents' behaviour. Studies of mind reading should control for submentalising because cognitive science has shown that domain-general mechanisms of attention, learning and memory are activated in adult humans whenever patterns recur in a complex stimulus array (Box 1). Therefore, to consider the possibility that apes are submentalising is not to belittle them, but to ask in what ways they are similar to humans. Unless one needs to discuss behaviour [9], or to catch a Hollywood spy, submentalising may be the smart option.

References

1. Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and brain sciences*, *1*(04), 515-526.

2. Heyes, C. M. (1998). Theory of mind in nonhuman primates. *Behavioral and Brain Sciences*, *21*(01), 101-114.

3. Heyes, C. (2015). Animal mindreading: what's the problem? *Psychonomic bulletin & review*, *22*(2), 313-327.

4. 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.

5. Heyes, C. (2014). Submentalizing: I am not really reading your mind. *Perspectives on Psychological Science*, *9*(2), 131-143.

6. von Uexküll, J. (1934). A stroll through the worlds of animals and men (trans: Claire H. Schiller). *Instinctive behavior, ed. Claire H. Schiller*, 5-80.

7. Tinkelpaugh, O. L. (1932). Multiple delayed reaction with chimpanzee and monkeys. *Journal of Comparative Psychology*, *13*, 207-243.

8. Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive psychology*, *36*(1), 28-71.

9. Heyes, C. M., & Frith, C. D. (2014). The cultural evolution of mind reading. *Science*, *344*(6190), 1243091.

10. Heyes, C. (2014). False belief in infancy: a fresh look. *Developmental Science*, *17*(5), 647-659.

Box 1. Testing for submentalising.

Experiments on mind reading can be controlled for submentalising in several ways [3, 5, 10]. A particularly effective strategy uses inanimate control stimuli of the kind shown in Figure 1. In these examples, based on frames in the 'belief induction' (A) and 'false belief' (B) movies presented by Krupenye and colleagues, the green actor and KK have been replaced by coloured shapes. If the apes tested by Krupenye and colleagues had shown the same eye-movement behaviour after viewing a movie that featured colors, shapes and movements without human actors, it would suggest their eye movements were due to submentalising.

It is also a good idea to use within-subject designs, involving transfer tests, to pit mentalizing against submentalising hypotheses [2, 3]. When an animal behaves in the same way across conditions that support the same mental state attribution, but involve stimuli with different low-level features, it is more likely that their behaviour is being guided by the high-level rather than low-level properties of the stimuli. Krupenye and colleagues put between-subjects analysis of their data in the foreground. However, to their credit, they also reported within-subjects analyses which did not support their mentalizing hypothesis. The data showed that the apes did not show transfer across stimulus sequences; performance in Experiments 1 and 2 was not correlated, and the "follow-up" did not replicate the results of Experiment 1. Therefore, in principle a different submentalising mechanism could have been responsible for the behaviour of each group of apes in each experiment.

Figure 1. Inanimate control stimuli can distinguish mentalising from submentalising.