Designer receptor inhibition suggests mechanism for monkey Theory of Mind

Robert R. Hampton¹

Accepted: 13 November 2020 / Published online: 24 November 2020
© The Psychonomic Society, Inc. 2020

Summary

Theory of Mind (ToM) is a critical component of human social cognition that has not been found reliably in other primates, especially monkeys. Hayashi et al. (*Cell Reports*, 30, 4433–4444, 2020) used newly developed behavioral techniques to detect evidence for ToM in Japanese macaque monkeys and employed sophisticated neurobiological manipulations to implicate the medial prefrontal cortex in this capacity.

Keywords Theory of mind · False belief · Anticipatory looking · Primate · Monkey · Cognition

Humans routinely explain and understand the behavior of others in terms of beliefs, a capacity often called Theory of Mind (ToM). We explain Jon looking for his keys on his desk by stating that he believes they might be there. This is not “mind reading.” It is done on the basis of observations or inferences about the past experience and current behavior of the individual we are trying to understand (Povinelli & Vonk, 2003). Because ToM appears to play such a critical role in human social behavior, and because deficits in ToM seem to underlie various deficits in social behavior, the study of ToM in nonhuman animals is of great interest to both students of cognitive evolution and investigators striving to understand the neural basis of normal and pathological social behavior. A recent paper by Hayashi et al. (2020) advances work on both fronts.

Performance in so-called false-belief tasks has become a standard for assessing ToM because these tasks dissociate behavior controlled by the actual state of the world from that controlled by beliefs. The core feature of these tasks is arrangement of circumstances such that an observer has the opportunity to learn that another individual has an incorrect belief about the world. For example, in “change-of-location” tasks, the observer sees “Joni” watch food being hidden in location A. Joni leaves, and while she is away, the

experimenter or some other agent moves the food to location B. The observer thus knows that the food is in location B, but could also know that Joni still *believes* that the food is in location A because Joni did not see the food moved. The critical test is to determine where the observer expects Joni to look for the food when she returns. If the observer has ToM, they should expect Joni to look in location A, because that is where Joni believes the food is. An observer who passes this test appears to understand that it is Joni’s belief, rather than the actual state of the world, that controls her behavior.

It is perhaps ironic that while the concept of ToM was first proposed with respect to chimpanzees (Premack & Woodruff, 1978), there has been sustained debate ever since regarding the existence of ToM in nonhuman primates, while evidence for this capacity in humans has been widely accepted. Three aspects of the procedure used by Hayashi et al. (2020) may account for the striking positive evidence for ToM in monkeys they report. First, they used “engaging social stimuli” to implement their change-of-location task. For example, humans appeared to compete over a toy or hide from one another. Arguably, these engaging stimuli ensure monkeys attended to the stimulus videos. Second, the dependent measure used was anticipatory looking, rather than either the explicit choice measures or expectancy violation measures used in most prior work. Avoiding explicit choice admits the possibility that monkey ToM may be implicit. Anticipatory looking may be superior to expectancy violation because it is a cleaner measure of what the observer monkey expects “Joni” to do, whereas the observer could be surprised by Joni’s behavior for a variety of reasons that would be confounded in a study

✉ Robert R. Hampton
robert.hampton@emory.edu

¹ Department of Psychology and Yerkes National Primate Research Center, Emory University, Atlanta, GA, USA

using expectancy violation. Third, Hayashi et al. used the “Southgate” method pioneered in work with children (Southgate, Senju, & Csibra, 2007). Rather than leaving the hidden item in location B, the observer sees the item removed from the experimental scene entirely. Thus, when Joni returns to the experimental scene, the observer is not presented with information that conflicts with Joni’s false belief that the target item is in location A. There is currently no target item anywhere in the experimental scene. This may provide a more sensitive measure of false belief understanding because the observer’s understanding of false belief is the only systematic influence on where they should expect Joni to look (see Horschler, MacLean, & Santos, 2020, for a thorough, insightful, and cautionary review of related issues).

Whether it is the engaging social content, anticipatory looking, the “Southgate” method, or some combination of these that was critical, Hayashi et al. (2020) reported evidence clearly in support of some understanding of false belief and ToM by Japanese macaque monkeys (*Macaca fuscata*). Across three different video scenarios involving a change-of-location to generate false belief, observer monkeys looked first and longer at the location where “Joni” had seen the target item hidden, apparently anticipating her choice behavior when she returned to the experimental area. Observer monkeys showed this bias even though “Joni” was positioned equidistant between the two locations. This is an exciting finding and raises at least three general classes of questions that deserve the attention of students of cognitive evolution. First, does this indeed mean that the capacity for ToM first evolved at least 30+ million years ago when a common ancestor of old-world monkeys and apes lived? Second, is it true that evidence for understanding false belief is only evident in implicit tests? If so, what does it mean functionally for this sophisticated social cognition to occur only implicitly in monkeys? How did this capacity become explicit in humans? Third, how widespread is ToM and understanding of false belief as measured by this new method? Perhaps it will be more widespread, even outside primates, than many of us expected.

As if collecting this behavioral evidence for ToM in monkeys was not important enough, Hayashi et al. (2020) went on to use the sophisticated neurobiological manipulation of DREADDs (Designer Receptors Exclusively Activated by Designer Drugs) to leverage their behavioral discovery for the neuroscience of social cognition. They tested whether the medial prefrontal cortex (MPFC), part of a network identified through functional magnetic resonance imaging (fMRI) to be active during human social cognition, is causal in ToM.

They used a lentivirus vector to infect neurons in the MPFC with receptors for clozapine N-oxide (CNO), a designer drug intended to bind selectively and uniquely to these new receptors. CNO administered by peripheral injection later inhibited the activity of neurons infected by the lentivirus vector, a fact confirmed electrophysiologically.

Injection of CNO in monkeys with the designer CNO receptor also reduced the extent to which monkeys showed evidence of understanding false belief in the change-of-location task, indicating a critical role for medial prefrontal cortex in ToM. These treatments appeared not to affect attention to the videos, or gaze tracking generally. Control injections of CNO in monkeys lacking the designer receptor indicated that CNO alone did not account for the change in behavior. Thus, Hayashi et al. (2020) have provided us with a potentially powerful new primate model system for causal studies of the neurobiology of ToM. The effects of the DREADDs manipulation on performance in the false-belief test were not as robust as the initial behavioral findings, and therefore replications should be attempted. This work is a dramatic combination of sophisticated and complex behavioral and neurobiological work and will be challenging for other teams to implement, but important to attempt.

While we don’t yet know with certainty whether it is a true or false belief that macaque monkeys have ToM, we should be looking with anticipatory interest for new studies using the techniques described by Hayashi et al. (2020), and we should be prepared to have our expectations violated.

References

- Hayashi, A., Kawasaki, E., Minamimoto, K., ... Hasegawa. (2020). Macaques Exhibit Implicit Gaze Bias Anticipating Others' False-Belief-Driven Actions via Medial Prefrontal Cortex. *Cell Reports*, 30(13), 4433. <https://doi.org/10.1016/j.celrep.2020.03.013>
- Horschler, M., & Santos. (2020). Do Non-Human Primates Really Represent Others' Beliefs? *Trends In Cognitive Sciences*, 24(8), 594-605. <https://doi.org/10.1016/j.tics.2020.05.009>
- Povinelli, & Vonk (2003). Chimpanzee minds: suspiciously human? *Trends In Cognitive Sciences*, 7(4), 157-160.
- Premack, & Woodruff (1978). Does the chimpanzee have a theory of mind? *Behav. Brain Sciences*, 4, 515-526.
- Southgate, S., & Csibra. (2007). Action anticipation through attribution of false belief by 2-year-olds. *Psychological Science*, 18(7), 587-592. <https://doi.org/10.1111/j.1467-9280.2007.01944.x>

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.