



Reply to Royka and Santos: Representing ignorance, and the evolution of theory of mind

Luke A. Townrow^{a,1}  and Christopher Krupenye^{a,1} 

In a recent study (1), we showed that bonobos pointed more often and more quickly toward a hidden food reward when their cooperative human partner was ignorant of the food's location than when he was knowledgeable. One goal of this work was to test whether apes can represent mental states that may be decoupled from reality, like ignorance, or only knowledge-like states (2). In a thoughtful letter, Royka and Santos (3) argue that our data remain consistent with both interpretations: Bonobos might show elevated pointing because they desire the food and, without a representation of their partner's ignorance, do not know whether their partner will otherwise obtain it for them. To provide more unambiguous evidence for ignorance attribution, Royka and Santos propose investigating whether primates make positive predictions about the actions of ignorant agents (e.g., whether primates will expect an ignorant agent to seek information or be surprised by an ignorant agent's improbable discovery of hidden food).

It is possible to interpret our results as at least partial evidence for such positive predictions: According to this view, bonobos predict that their ignorant partner is unlikely to find the food and therefore point to request that he search in a particular location. We also find this interpretation to be maximally consistent with the broader literature. While there is little unambiguous evidence for ignorance attribution in the literature, several studies now provide evidence that apes can track other decoupled states, like beliefs (e.g., refs. 4–7). However, we agree with Royka and Santos that it is not possible to completely rule out their interpretation by appealing to our experiment alone. We are also excited by and fully endorse their suggestion that researchers test whether primates make positive predictions of ignorant individuals, and more generally look forward to future experimental tasks that test against one another the divergent predictions of

viable theoretical accounts. Such tasks hold the promise of isolating the cognitive representations that enable animals to predict and understand the behavior and minds of other individuals.

One exciting possibility is that these tasks reveal important differences between species, suggesting different theories of mind across taxa, and potentially providing evidence for an evolutionary elaboration of theory of mind (8). For example, perhaps extant monkeys, and our extinct 30 MYO common ancestor, represent more constrained mental states than apes (2, but see ref. 9). Apes, and our more recent common ancestor with them, may have evolved a richer capacity to track knowledge, ignorance, and simple kinds of beliefs. Humans may have evolved capacities that are richer still. Perhaps we possess unique computational resources for tracking more complex beliefs, as in the Gettier case (10), or those involving identity, level II perspective-taking, or greater recursion (8). We are excited by the prospect that new experimental paradigms, like our own and those of Royka et al. (11), will deliver answers to these age-old questions about the evolution of the human mind, what makes it unique, and the forms of thought that can exist in the absence of language.

Author affiliations: ^aDepartment of Psychological and Brain Sciences, Johns Hopkins University, Baltimore, MD 21218

Author contributions: L.A.T. and C.K. designed research and wrote the paper.

The authors declare no competing interest.

Copyright © 2025 the Author(s). Published by PNAS. This article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: ltownro1@jh.edu or krupenye@jh.edu.

Published May 9, 2025.

1. L. A. Townrow, C. Krupenye, Bonobos point more for ignorant than knowledgeable social partners. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2412450122 (2025).
2. A. Martin, L. R. Santos, What cognitive representations support primate theory of mind? *Trends Cogn. Sci.* **20**, 375–382 (2016).
3. A. Royka, L. R. Santos, Informing does not require attributing ignorance. *Proc. Natl. Acad. Sci. U.S.A.* **122**, e2504155122 (2025).
4. C. Krupenye, F. Kano, S. Hirata, J. Call, M. Tomasello, Great apes anticipate that other individuals will act according to false beliefs. *Science* **354**, 110–114 (2016).
5. F. Kano, C. Krupenye, S. Hirata, M. Tomonaga, J. Call, Great apes use self-experience to anticipate an agent's action in a false-belief test. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 20904–20909 (2019).
6. D. Buttelmann, F. Buttelmann, M. Carpenter, J. Call, M. Tomasello, Great apes distinguish true from false beliefs in an interactive helping task. *PLoS One* **12**, e0173793 (2017).
7. R. W. Lurz, C. Krachun, M. C. Marengo, W. D. Hopkins, Do chimpanzees predict others' behavior by simulating their beliefs. *Anim. Behav. Cogn.* **9**, 153–175 (2022).
8. C. Krupenye, "The evolution of mentalizing in humans and other primates" in *The Neural Basis of Mentalizing*, M. Gilead, K. N. Ochsner, Eds. (Springer International Publishing, Cham, 2021), pp. 107–129.
9. T. Hayashi et al., Macaques exhibit implicit gaze bias anticipating others' false-belief-driven actions via medial prefrontal cortex. *Cell Rep.* **30**, 4433–4444 (2020).
10. K. Durdevic, C. Krupenye, Representing knowledge, belief, and everything in between: Representational complexity in humans and other apes. *Behav. Brain Sci.* **44**, e150 (2021).
11. A. L. Royka, D. J. Horschler, W. Bargmann, L. Santos, "Probing nonhuman primate errors on false belief tasks to explore the evolutionary roots of theory of mind" in *Proceedings of the 46th Annual Conference of the Cognitive Science Society*, L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, E. Hazeltine, Eds. (Cognitive Science Society, Rotterdam, The Netherlands 2024), pp. 1661.