

Informing does not require attributing ignorance

Amanda Royka^{a,1} and Laurie R. Santos^a

Townrow et al. (1) provide elegant evidence that bonobos act to shape others' behaviors. Their study also reveals this capacity's flexibility across nonhuman great apes: They shape others' future behaviors in different contexts (1–3) and do so using both naturalistic (2, 3) and learned behaviors (1).

Unfortunately, we are not as convinced that ref. 1 provides evidence that bonobos represent others' ignorance. To illustrate why, imagine the following scenario: A witch has chained you to the wall of a dungeon and placed the key outside a nearby window. Now imagine that a stranger walks by the window. You have no idea what this stranger knows or does not know, yet you can easily simulate that banging on the window and pointing to the key could cause the stranger to turn and act on the key, which could lead to your escape. We argue that subjects tested in ref. 1 could succeed on the "ignorance" trials in a similar way: They begin with *no representation* about the experimenter's mental states.

If, as we argue, bonobos in ref. 1 could perform well either by representing the experimenter's ignorance (as Townrow et al. argue; Fig. 1*B*) or by representing *no information* about the agent's mental states (Fig. 1*C*), how can we differentiate between these two explanations? One possibility would be to show that bonobos are able to make *positive predictions* about how an ignorant agent will behave. Let us return to the dungeon scenario. Let us say that the person walking by the window was not a stranger, but your friend—someone you had reason to suspect was totally ignorant of your situation. In that case, you would probably have *positive predictions* about his behavior based on that state of ignorance. You might predict, for example, that your friend would *information-seek*: He might ask you what was going on and how you got locked inside. He might also look for *new* information, such as trying to determine how the dungeon lock

Author affiliations: ^aDepartment of Psychology, Yale University, New Haven, CT 06510

Author contributions: A.R. and L.R.S. 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: amanda.royka@yale.edu. Published May 9, 2025.



Fig. 1. Different ways that bonobos tested in Townrow et al. (1) might represent the experimenter's mental state. Critically, representing different mental states would change the subject's predictions about what the experimenter might do next, which we represent here as different probability distributions over his possible next action. (*A*) In the knowledgeable trials, subjects know that the experimenter has seen where the food is hidden so they represent him as aware of the location of the food and predict that he is very likely to act on that location. As such, subjects recognize that, in those trials, they do not need to point. By contrast, in the ignorance trials, there are two possibilities that could cause the subjects to point to the correct cup. (*B*) Subjects could succeed in the ignorance trials because they represent that the experimenter will act on that location (i.e., they think all cups are equally likely to be searched) and thus point to the correct cup to increase the likelihood that the experimenter will act on that location. (C) Alternatively, bonobos may not have any representation of the experimenter would behave. Under this scenario, all possible actions have equal probability and thus subjects would have no positive predictions about how the experimenter would behave. Therefore, subjects also point to the correct cup in order to increase the likelihood that the experimenter will act on that location.

works. Moreover, you would also be surprised if your ignorant friend made an improbable guess about this strange situation (e.g., "Wait, was it a witch that locked you up?"). In all these cases, you would be using your representation of that friend's ignorance to make positive predictions about his next actions—predictions that would be different than if you simply were not sure what someone knew about the scenario.

To truly show that bonobos understand ignorance, researchers must demonstrate that they make *positive*

predictions about how an ignorant agent will behave (4). For example, do bonobos predict that an ignorant agent will search for information that they do not yet know? Would bonobos be surprised if an ignorant agent correctly made an improbable guess on their first try (e.g., guessing the correct hiding location of a reward placed in one of 50 possible locations)? Without such findings, we worry that while ref. 1 provides insight into bonobos' communicative prowess, the results fall short of revealing their underlying representational capacities.

- 2. C. Crockford, R. M. Wittig, R. Mundry, K. Zuberbühler, Wild chimpanzees inform ignorant group members of danger. Curr. Biol. 22, 142–146 (2012).
- 3. C. Crockford, R. M. Wittig, K. Zuberbühler, Vocalizing in chimpanzees is influenced by social-cognitive processes. Sci. Adv. 3, e1701742 (2017).
- 4. A. Martin, L. R. Santos, What cognitive representations support primate theory of mind? *Trends Cogn. Sci.* 20, 375–382 (2016).

^{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).