Multiple representational theories explain non-human primate perspective-taking: Evidence from computational modeling

Daniel Horschler
Yale University, New Haven, Connecticut, United States

Marlene Berke
Yale University, New Haven, Connecticut, United States

Laurie Santos
Yale University, New Haven, Connecticut, United States

Julian Jara-Ettinger
Yale University, New Haven, Connecticut, United States

Abstract

Humans' ability to attribute mental states to agents has been hypothesized to underpin our unique social behaviors. However, questions remain about the extent to which our representational Theory of Mind (ToM) is shared with non-human primates (NHPs). Here, we present a set of computational models each built to formalize a different representational theory of a foundational ToM component—understanding what others can see—and compare each model's performance to that of NHPs across a range of previously published perspective-taking experiments. Our results show that multiple competing theories can account for NHPs' perspective-taking abilities, including both human-like ToM and less complex mentalistic theories, but not simpler, non-mentalistic theories. This work supports the idea that NHPs may reason about others' mental states when assessing their visual perspectives, and provides promising avenues for future work using computational modeling to determine if and how NHPs represent more complex mental states (e.g., ignorance, belief).

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*.