

- Surtees, A. D., & Apperly, I. A. (2012). Egocentrism and automatic perspective taking in children and adults. *Child Development*, 83(2), 452–460.
- Surtees, A. D. R., Butterfill, S. A., & Apperly, I. A. (2012). Direct and indirect measures of level-2 perspective-taking in children and adults. *British Journal of Developmental Psychology*, 30, 75–86. doi: 10.1111/j.2044-835X.2011.02063.x.
- Tahiroglu, D., Moses, L. J., Carlson, S. M., Mahy, C. E., Olofson, E. L., & Sabbagh, M. A. (2014). The children's social understanding scale: Construction and validation of a parent-report measure for assessing individual differences in children's theories of mind. *Developmental Psychology*, 50(11), 2485–2497.
- Tan, J., & Harris, P. L. (1991). Autistic children understand seeing and wanting. *Development and Psychopathology*, 3(2), 163–174.
- Tardif, T., & Wellman, H. M. (2000). Acquisition of mental state language in Mandarin-and Cantonese-speaking children. *Developmental Psychology*, 36(1), 25.
- Tomasello, M. (1999). *The cultural origins of human cognition*. Harvard University Press.
- Tomasello, M. (2018). How children come to understand false beliefs: A shared intentionality account. *Proceedings of the National Academy of Sciences*, 115(34), 8491–8498.
- Tomasello, M., Call, J., & Hare, B. (2003). Chimpanzees understand psychological states – the question is which ones and to what extent. *Trends in Cognitive Sciences*, 7(4), 153–156.
- Tomasello, M., Kruger, A., & Ratner, H. (1993). Cultural learning. *Behavioral and Brain Sciences*, 16, 495–552.
- Träuble, B., Marinović, V., & Pauen, S. (2010). Early theory of mind competencies: Do infants understand others' beliefs? *Infancy*, 15(4), 434–444.
- Trick, L. M., & Pylyshyn, Z. W. (1994). Why are small and large numbers enumerated differently? A limited-capacity preattentive stage in vision. *Psychological Review*, 101(1), 80–102. <http://dx.doi.org/10.1037/0033-295X.101.1.80>.
- Tummelshammer, K. S., Wu, R., Sobel, D. M., & Kirkham, N. Z. (2014). Infants track the reliability of potential informants. *Psychological Science*, 25(9), 1730–1738.
- Turri, J. (2015a). Evidence of factive norms of belief and decision. *Synthese*, 192, 4009–4030. doi: 10.1007/s11229-015-0727-z.
- Turri, J. (2015b). Knowledge and the norm of assertion: A simple test. *Synthese*, 192(2), 385–392. doi: 10.1007/s11229-014-0573-4.
- Turri, J. (2016a). Knowledge attributions and behavioral predictions. *Cognitive Science*, 41(8), 2253–2261.
- Turri, J., & Buckwalter, W. (2017). Descartes's schism, Locke's reunion: Completing the pragmatic turn in epistemology. *American Philosophical Quarterly*, 54(1), 25–46.
- Turri, J., Buckwalter, W., & Rose, D. (2016). Actionability judgments cause knowledge judgments. *Thought: A Journal of Philosophy*, 5(3), 212–222. <http://doi.org/10.1002/tht3.213>.
- Turri, J., Friedman, O., & Keefner, A. (2017). Knowledge central: A central role for knowledge attributions in social evaluations. *Quarterly Journal of Experimental Psychology*, 70(3), 504–515.
- van der Wel, R. P., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition*, 130(1), 128–133.
- Vouloumanos, A., Martin, A., & Onishi, K. H. (2014). Do 6-month-olds understand that speech can communicate? *Developmental Science*, 17(6), 872–879.
- Wellman, H. M. (2014). *Making minds: How theory of mind develops*. Oxford University Press.
- Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development*, 72, 655–684.
- Wellman, H. M., Fang, F., Liu, D., Zhu, L., & Liu, G. (2006). Scaling of theory-of-mind understandings in Chinese children. *Psychological Science*, 17(12), 1075–1081.
- Wellman, H. M., & Liu, D. (2004). Scaling of theory-of-mind tasks. *Child Development*, 75(2), 523–541.
- Whiten, A. (2013). Humans are not alone in computing how others see the world. *Animal Behaviour*, 86(2), 213–221.
- Williamson, T. (2000). *Knowledge and its limits*. Oxford University Press. doi:10.1093/019925656X.001.0001.
- Woolley, J. D., & Wellman, H. M. (1993). Origin and truth: Young children's understanding of imaginary mental representations. *Child Development*, 64(1), 1–17.
- Wynn, K. (1992). Addition and subtraction by human infants. *Nature*, 358(6389), 749–750.
- Xu, F., & Spelke, E. S. (2000). Large number discrimination in 6-month-old infants. *Cognition*, 74(1), B1–B11.
- Yirmiya, N., Sigman, M., & Zacks, D. (1994). Perceptual perspective-taking and seriation abilities in high-functioning children with autism. *Developmental Psychopathology*, 6, 263–272. doi: 10.1017/S0954579400004570.
- Zmyj, N., Buttelmann, D., Carpenter, M., & Daum, M. M. (2010). The reliability of a model influences 14-month-olds' imitation. *Journal of Experimental Child Psychology*, 106(4), 208–220.

Open Peer Commentary

Beyond knowledge versus belief: The contents of mental-state representations and their underlying computations

Mika Asaba , Aaron Chuey, and Hyowon Gweon 

Department of Psychology, Stanford University, Stanford, CA 94305, USA.
masaba@stanford.edu; chuey@stanford.edu; hyo@stanford.edu
<https://sll.stanford.edu>

doi:10.1017/S0140525X20001879, e141

Abstract

Moving beyond distinguishing knowledge and beliefs, we propose two lines of inquiry for the next generation of theory of mind (ToM) research: (1) characterizing the contents of different mental-state representations and (2) formalizing the computations that generate such contents. Studying how children reason about what others think of the self provides an illuminating window into the richness and flexibility of human social cognition.

We agree with Phillips et al. that examining a greater variety of epistemic states will enrich our understanding of the origins and development of mentalizing capacities, and appreciate their distinction between knowledge and belief. Importantly, however, categorizing these mental states and asking at what age children can represent them are just the first steps toward characterizing the richness and flexibility of our social-cognitive capacities. Looking back on over a decade of research attempting to identify the earliest signatures of belief attribution (e.g., Onishi & Baillargeon, 2005; Surian, Caldi, & Sperber, 2007; but see also Dörrenberg, Rakoczy, & Liszkowski, 2018; Powell, Hobbs, Bardis, Carey, & Saxe, 2018), we caution against yet another arms race to determine which representation is “more basic” or “more critical” to social learning. Instead, as cognitive scientists, we find ourselves asking: *How* are these mental-state representations cognitively distinct from one another, and what cognitive mechanisms support these representations?

Imagine your roommate is watching as you try to open a jar that just won't budge. You might say that your roommate *knows* you failed to open the jar, *believes* the jar is difficult to open, or even *thinks* that you are too weak to open it. Although you would be comfortable using these words – know, believe, think – to describe your roommate's mental states, their contents differ from what is typically studied in the theory of mind (ToM) literature. Rather than reflecting verifiable external states of the world (e.g., the location of the jar), these mental states concern outcomes of intentional, goal-directed actions (e.g., failure to open a jar) and outputs of additional inferences based on observed action–outcome relationships, such as subjective evaluations about abstract qualities of objects (e.g., the jar may be difficult to open for some but not others) and agents, including oneself (e.g., you may be weak compared to some but not others).

To understand how our mind flexibly generates, represents, and attributes these mental states, we need to move beyond traditional concepts and empirical paradigms that have dominated the past

few decades of research on ToM. To this end, young children's reasoning about how others represent and infer abstract qualities of people (including the self) can provide a particularly illuminating window into the richness and flexibility of human social cognition. Our recent study finds that children, by 4 years of age, are already capable of reasoning about what others think of them after observing their own failures or successes (Asaba & Gweon, 2018; Asaba & Gweon, 2021). Looking forward, we propose focusing our efforts on two related lines of inquiry: (1) characterizing the *contents* of mental states that children (and nonhuman primates) can represent and (2) understanding and formalizing the *computations* (i.e., inferential processes) that give rise to such representations.

1. Contents of mental states

If your roommate observed your failure to open the jar countless times every day, you may intuitively feel that your roommate "knows" you cannot open the jar. Similarly, we might use the word "know" to describe one's various attitudes toward someone else (e.g., "Sally knows that Ann is generous, funny, and competent"), especially when we suspect one has strong evidence about these qualities. However, these are inherently subjective evaluations that do not have objective, verifiable criteria for determining their truth value; they can only be expressed as the degree to which one "believes" X is true, rather than as a Boolean value (i.e., either true or false). Although people often describe these mental states using knowledge-laden language, these contents go beyond the scope of what Phillips et al. would consider as knowledge.

Critically however, the content of these representations also differ from the content of beliefs that are typically studied in the ToM literature. Rather than observable states of the physical world that are verifiable via perception (e.g., "Sally knows her toy is in the box" or "Sally thinks her toy is in the box"), these belief states concern abstract properties of agents that must be inferred from an agent's behaviors or other social sources of information (e.g., others' evaluative feedback or testimony, such as "Ann is very generous"). Beyond distinguishing knowledge versus belief, we need more research on *why* children find some mental-state contents harder to attribute than others.

2. Underlying computations

Relatedly, the process by which we attribute mental states about internal qualities of agents may involve more complex computations than those concerning external states of the world. When an agent has direct perceptual access to a world-state, there is a one-to-one correspondence between what they see and what they represent. When the agent loses perceptual access while the world-state changes, the agent is rendered ignorant (i.e., Sally does not know where her toy is) or mistaken (i.e., Sally falsely believes that her toy is in the box). Although understanding the relationship between an agent's perception and their resulting epistemic state is already an impressive feat, representing others' beliefs about internal qualities is even more so; these representations cannot be derived from perceptual access alone, and require further inferences based on an intuitive understanding of how observations of an agent's goal-directed actions give rise to representations of the agent's abstract qualities.

Much of the prior literature on ToM development has studied how children represent others' ignorance or false beliefs that are decoupled from reality. However, these representations reflect only a fraction of the mental states we encounter in our everyday

social interactions. When Sally observes Ann bake a delicious cake, get a "D" on a math exam, or donate \$20 dollars to charity, what kind of mental states might children attribute to Sally? These representations could be about anyone, but they are especially powerful when they concern qualities of the self: Does Sally think I am good at baking? Terrible at math? Generous or stingy? Although we, as adults, naturally entertain these thoughts, more research is needed to understand how young children integrate their understanding of the physical and social world to attribute these nuanced mental states. Recent computational research has made major advances in formalizing the generative process by which an agent's observation gives rise to beliefs about the external world (Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017; Jara-Ettinger, Schulz, & Tenenbaum, 2020); these approaches can provide important insights here as well.

Ultimately, the ability to represent others' mental states is both a blessing and a curse. When directed at us, it inspires our motivation to learn and responsibility for our own actions; when gone awry, it dogs us with unnecessary worries about how others might evaluate us. Yet, for better or worse, we rely on these abilities to learn from others in a complex social world filled with competition, cooperation, and collaboration. The richness of human cultural knowledge comes from our ability to appreciate, evaluate, criticize, and communicate a host of abstract thoughts. By studying how the human mind supports these rich mental-state inferences, we can better understand how humans harness these capacities to learn from others and help others learn.

References

- Asaba, M., & Gweon, H. (2018). Look, I can do it! Young children forego opportunities to teach others to demonstrate their own competence. In C. Kalish, M. Rau, J. Zhu, & T. Rogers (Eds.), *Proceedings of the 40th annual conference of the cognitive science society* (pp. 106–111). Cognitive Science Society.
- Asaba, M., & Gweon, H. (2021). Young children rationally revise and maintain what others think of them. *PsyArxiv*. Preprint doi: [10.31234/osf.io/yxhv5](https://doi.org/10.31234/osf.io/yxhv5).
- Baker, C. L., Jara-Ettinger, J., Saxe, R., & Tenenbaum, J. B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1, 1–10. doi: [10.1038/s41562-017-0064](https://doi.org/10.1038/s41562-017-0064).
- Dörrenberg, S., Rakoczy, H., & Liszkowski, U. (2018). How (not) to measure infant theory of mind: Testing the replicability and validity of four non-verbal measures. *Cognitive Development*, 46, 12–30. doi: [10.1016/j.cogdev.2018.01.001](https://doi.org/10.1016/j.cogdev.2018.01.001).
- Jara-Ettinger, J., Schulz, L. E., & Tenenbaum, J. B. (2020). The naïve utility calculus as a unified, quantitative framework for action understanding. *Cognitive Psychology*, 123, 101334. doi: [10.1016/j.cogpsych.2020.101334](https://doi.org/10.1016/j.cogpsych.2020.101334).
- Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, 308(5719), 255–258. doi: [10.1126/science.1107621](https://doi.org/10.1126/science.1107621).
- Powell, L. J., Hobbs, K., Bardis, A., Carey, S., & Saxe, R. (2018). Replications of implicit theory of mind tasks with varying representational demands. *Cognitive Development*, 46, 40–50. doi: [10.1016/j.cogdev.2017.10.004](https://doi.org/10.1016/j.cogdev.2017.10.004).
- Surian, L., Caldi, S., & Sperber, D. (2007). Attribution of beliefs by 13-month-old infants. *Psychological Science*, 18(7), 580–586. doi: [10.1111/j.1467-9280.2007.01943.x](https://doi.org/10.1111/j.1467-9280.2007.01943.x).

Infants actively seek and transmit knowledge via communication

Marina Bazhydai^a and Paul L. Harris^b

^aPsychology Department, Fylde College, Lancaster University, Lancaster LA1 4YW, UK and ^bHarvard Graduate School of Education, Cambridge, MA 02138, USA.

m.bazhydai@lancaster.ac.uk

paul_harris@gse.harvard.edu

doi:10.1017/S0140525X20001405, e142