Cognitive Architecture of Belief Reasoning in Children and Adults: A Primer on the Two-Systems Account

Jason Low,[1] Ian A. Apperly,[2] Stephen A. Butterfill,[3] and Hannes Rakoczy[4]

[1]*Victoria University of Wellington,* [2]*University of Birmingham,* [3]*University of Warwick, and* [4]*University of Göttingen*

ABSTRACT—***Characterizing the cognitive architecture of human mindreading forces us to address two puzzles in people's attributions of belief: Why children show inconsistent expectations about others' belief-based actions, and why adults' reasoning about belief is sometimes automatic and sometimes not. The seemingly puzzling data suggest that humans have many mindreading systems that use different models of mental representations. The efficient system is shared by infants, children, and adults, and uses a minimal model of the mind, which enables belief-like states to be tracked. The flexible system develops late and uses a canonical model, which incorporates propositional attitudes. A given model's operation has signature limits that produce performance contrasts, in children as well as adults, between certain types of mindreading tasks.***

KEYWORDS—*belief reasoning; perspective taking; two-systems theory; development; adults*

Our mindreading ability helps us reason about how beliefs might influence people's actions, interpersonal communications, and conduct. Research on people's attribution of beliefs reveals two puzzles: Children show an apparently contradictory pattern of success and failure in their responses to scenarios involving others' belief-based actions, and reasoning about belief is both nonautomatic and automatic. To solve these puzzles, we highlight cognitive studies of children and adults to introduce an approach to the architecture of mindreading in which people can be of at least two minds about the ways in which others' beliefs cause and rationalize behavior (1, 2). We discuss how signature limits on low-level processes make it possible to differentiate efficient and flexible instances of mindreading. We then evaluate a contrasting account suggesting that human beings have a unitary, abstract psychological reasoning system from early in life.

Jason Low, Victoria University of Wellington; Ian A. Apperly, University of Birmingham; Stephen A. Butterfill, University of Warwick; Hannes Rakoczy, University of Göttingen.

Correspondence concerning this article should be addressed to Jason Low, School of Psychology, Victoria University of Wellington, P.O. Box 600, Wellington, New Zealand 6140; e-mail: Jason.Low@vuw.ac.nz.

DOI: 10.1111/cdep.12183

PUZZLES IN PEOPLE'S ATTRIBUTION OF BELIEF

Puzzle 1: Infants Pass False-Belief Tasks, but 3-Year-Olds Fail

A common measure of the development of our mindreading ability is the false-belief task. In one study (3), preschoolers watched as Maxi witnessed a target placed at location X. In Maxi's absence, the target was moved to location Y, and children were asked to predict where Maxi would look for the target. Most 3-year-olds answered that Maxi would look in location Y, as if false belief were impossible; by contrast, many 4-year-olds said Maxi would look in location X, indicating that they recognized Maxi's false belief. The incorporation of belief into children's understanding of minds from about age 4 is a well-replicated and robust finding (4). Once children master verbal false-belief tasks, they do so systematically and coherently for many topics and in many formats of tasks. Four-year-olds' grasp of beliefs includes appreciating that beliefs are essentially aspectual, that is, beliefs represent a given object under some guises but not others. In another study (5), when 4-year-olds gave correct verbal answers to standard false-belief tasks, they began to understand that an agent, depending on how he or she represents something, can believe mistakenly that two objects are present when, in fact, only one exists.

These findings from explicit verbal tasks contrast with results from nonverbal measures. Although 3-year-olds' verbal predictions indicated that they reasoned as if false belief were impossible, their eye movements in the same situation indicated that they could track others' false beliefs (6–8). This dissociation is supported by violation-of-expectation studies contrasting looking times to scenarios that are either consistent or inconsistent with an agent's belief. In one study (9), 15-month-olds saw scenarios of an agent forming either a true or false belief about an object's location. The agent searched in one of two locations—one compatible with the belief or another that was not. Infants looked longer when the agent searched in the location that was incompatible with the belief (longer looking was interpreted as infants expecting agents to act according to their beliefs). In other studies (10), 7- to 18-month-olds tracked false beliefs about contents and types of objects, and tailored their helping and communication to others' false belief about the location of an object. Therefore, the first puzzle is, how can infants and toddlers be sensitive to others' false beliefs when responding in some ways while they treat false belief as impossible when responding to the same situation in other ways?

Puzzle 2: Reasoning About Belief Is Both Nonautomatic and Automatic

Studies of human adults also point to seemingly incompatible sets of findings regarding the automaticity of mindreading inferences. A mindreading process is automatic if it occurs regardless of its relevance to participants' tasks and motives. In one study (11), false beliefs were not ascribed automatically: Adults with no specific motive to attend to a character's beliefs responded more slowly to unpredictable questions about an agent's false belief of an object's whereabouts than to the matched control probes. The case for nonautomaticity is also supported by research showing that tracking beliefs frequently depends on resources related to attention and working memory in adults and, furthermore, that even merely holding in mind someone else's belief incurs significant processing costs (12).

However, the results of two studies point to a different conclusion. In one (13), a character's false belief could influence adults' visual attention regardless of the relevance of the belief to the tasks adults were assigned. Adults who were told to track a character's belief and adults who were told to track a ball's location fixated longer on an empty box when the character had a false belief that the ball was in that location than when the character had a true belief that the ball was not in that location. Mirroring findings from young children, in another study (14), the effects of calculating indirect beliefs were different from the effects of judgments about direct beliefs. Adults saw a ball and a cube disappear behind two screens. A bystander had a false belief, while participants had a true belief about the objects' locations. Participants who were instructed to move a computer mouse to reach the ball's location showed involuntary tracking of belief: They moved the mouse to the ball in a way that was skewed toward where the bystander falsely believed the ball to be. Deliberate inferences showed different effects: Participants who were told to track beliefs took longer to move the mouse when their beliefs differed from the bystander's (and their mouse movements were not skewed by the bystander's beliefs). Therefore, the second puzzle is, why is belief tracking sometimes but not always automatic?

THE TWO-SYSTEMS ACCOUNT

We can solve these two puzzles by supposing that mindreading architecture involves at least two systems for tracking mental states, with complementary tradeoffs between efficiency and flexibility—in much the way that some theories feature at least two systems for tracking numbers (15). The efficient mindreading system is evolutionarily and ontogenetically ancient, operates quickly, and is largely automatic and independent of central cognitive resources. In contrast, the flexible mindreading system develops late, operates slowly, and makes substantial demands on executive control processes. Advances in executive function and language help cultivate flexible attributions about others' psychological perspectives (12). Although the efficient system typically subserves responses that occur independently of a participant's task and motives (e.g., looking behavior on some tasks), the flexible system is recruited by tasks that require declarative expressions of or deliberation about beliefs.

The processes that drive the efficient system may be triggered by direct cues like an agent's line of sight so that rapid online mindreading may be supported in participants with limited resources to process information. Deploying the flexible system does not depend on the immediate availability of cues about what a target witnesses. Components of efficient mindreading may have cognitive costs and may place some demands on working memory, as indicated by findings suggesting that dual tasking may disrupt looking-time responses to false-belief tasks (16). The efficient system should remain relatively distinct from the more flexible system, although the systems might exchange information over development (7, 12, 17).

Efficient mindreading is distinct from flexible mindreading in terms of signature limits arising from the type of model of mental representations that the respective systems rely on. The flexible system uses a canonical model of mental representations, where belief is characterized as a propositional attitude, that is, a state whose content can be picked out with a clause featuring the word *that* (e.g., Lucy believes that the morning star is above the horizon). A canonical model considers the aspectuality of beliefs, so that although the morning star is the same as the evening star, Lucy's belief that the morning star is above the horizon is distinct from her belief that the evening star is there. Such flexible reasoning would support understanding mistakes in others' representations of identity in the numerical sense, such as when Lucy believes falsely that the morning star is not the evening star. In contrast, the efficient system uses a minimal model of mental representations in which psychological states including belief-like states are characterized as relational attitudes, that is, states whose contents can be distinguished using relations between objects and locations or other properties.

Belief-like states can serve as proxies for beliefs: In a limited but useful range of situations, ascriptions of beliefs and belief-like states lead to identical expectations about an agent's behavior. However, the contents of belief-like states are not as fine grained as the truth conditions of beliefs; they are not aspectual (i.e., they do not distinguish the guises under which objects and situations are represented; 2). If Lucy has a belief-like relational attitude to the morning star and its position above the horizon, and if the morning star is the evening star, then she has the same attitude concerning the evening star. Therefore, an efficient mindreading system displays a signature limit concerning the aspectuality of belief.

Much like ascribing belief, reasoning about perception also involves more than tracking someone's visual connection to an object; different visual experiences may represent the same thing in different ways. An efficient mindreading system that is set to track relational attitudes will also be ill equipped to process the aspectual nature of mental states generally. Therefore, the two-systems account predicts that the efficient system can cover Level I visual-perspective-taking tasks (tracking what is or is not perceptible from different perspectives) and simple false-belief tasks about the location of objects (in which the participant has to keep track of what the agent has or has not witnessed). However, this system cannot cover Level II visual perspective taking (representing the way someone sees an object) or ascribing false beliefs about numerical identity, giving rise to signature limits.

SIGNATURE LIMITS ON EFFICIENT MINDREADING

At least three sources of evidence show how the difference in representational capacities of the two mindreading systems reveals itself in differential patterns of performance. First, in studies of visual perspective taking, humans automatically track what is seen but not how something is seen. In one study (18), adults saw photographs in which an avatar looked at all the dots on a wall (his perspective was consistent with participants') or the avatar looked at a subset of the dots (his perspective was inconsistent with participants'). Adults were slower and more prone to errors in judging how many dots they could see when the avatar's perspective differed from theirs. Another study (19) confirmed that adults experienced interference from the avatar's perspective only if they believed he could see, suggesting that interference resulted from processing of the avatar's mental states, not merely from seeing the direction he faced (cf. 20). Thus, even when calculating what others see (a Level I perspective-taking scenario) is irrelevant to the task, adults automatically track how others encounter and register objects, which interferes with self-judgments.

Fitting with the two-systems account, the interference in Level I perspective-taking scenarios does not generalize to Level II perspective-taking scenarios, which concern how an agent represents an object. In other studies (21, 22), children and adults did not automatically show such interference effects when participants had to report on the appearance of numerals, such as 6 and 9, that look different depending on the angle at which they are viewed. Similar patterns have been found in experiments measuring adults' eye movements as they process a speaker's instruction. For example, in one study (23), listeners looked more at an object that was visible to both the speaker and themselves, regardless of their own knowledge about the existence of objects in a scene. However, listeners were ineffective at taking the speaker's perspective regarding how an object was seen: Listeners looked more at an object that the speaker could not plausibly have been referring to because of the speaker's ignorance of the true identity of the visually misleading object.

In terms of the second source of evidence, in yet another study (24), 3- and 4-year-olds and adults looked in the correct location with the usual age-related improvements in verbal predictions when construing an agent's false belief about the location of an object. However, the same participants looked in the wrong location when tracking how an agent's representation of identities would lead to a false belief that there were two objects when, in fact, there was only one object (see Figure 1, Column 1). The switch from a location to an identity task did not affect verbal responses; 4-year-olds and most adults provided accurate verbal predictions. However, participants experienced the different visual aspects of the deceptive object late in the sequence. Demands associated with revising and updating inferences about the agent's representation of identities might have impaired participants' looking responses. That said, in a separate study (25), adults still looked in the wrong location but gave correct verbal predictions when the test object revealed its dual aspect early in the sequence (see Figure 1, Column 2).

In terms of the third source of evidence, researchers uncovered complementary findings when measuring toddlers' helping behavior (26). An agent watched as an object that appeared to be one toy (e.g., a rabbit) was put into a box. The object was removed from the box, transformed into a different toy (e.g., a carrot), and returned to the box, while the agent was watching in the true-belief condition, but while the agent was away in the false-belief condition. Then the agent watched the transformed object move to another box. Finally, the agent struggled to open the original box. Children behaved comparably in the false- and true-belief conditions: Most toddlers opened the original box, apparently focused on goal-directed relations. It is not that toddlers failed to understand identity per se; 14-month-olds can disregard superficial features and sort by object identity (27). When the false-belief task was switched to one of pure location tracking (cf. 28), children differentiated true- and false-belief conditions; most opened the original box in the former condition and the second box in the latter. Going beyond looking behavior, it appears that toddlers' helping actions that operate in concert

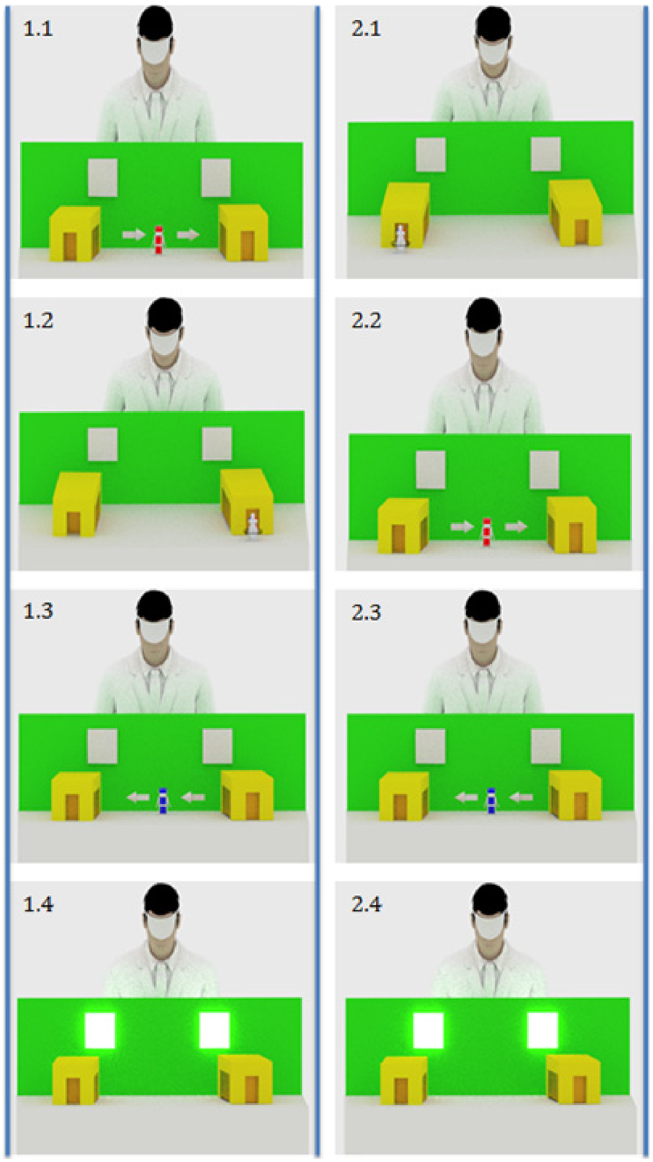

Figure 1. In the identity task (24; Column 1), the robot's red and blue aspects are revealed after it moved from the right-side Box A to the left-side Box B (Frame 1.2). Inside Box B, visible only to participants, the robot spun around to reveal its red and blue sides. Then the robot, with its blue aspect facing participants, moved back to Box A. If viewers represent object identities, they should anticipate that the agent falsely believes that another (blue) robot is inside Box B. The agent (e.g., having a blue-color preference) would have reason to reach into Box B. If participants tracked object registrations, then the robot is inside Box A and the agent should search there. In a modified version of the task (25; Column 2), dual identity was revealed inside Box A before the robot's first movement (Frame 2.1). In both versions, participants looked incorrectly (at Box A) with age-related increases in accuracy of verbal predictions.

with their early mindreading ability are also subject to characteristic signature limits.

In summary, the findings that adults' processing of how something is seen is not automatic by default and, furthermore, that children's anticipatory eye movements and helping responses appear insensitive to whether a protagonist has a false belief about the identity of an object lead us to conclude that cases of subjective mental representations are beyond the scope of the efficient mindreading system. Because such cases require reasoning based on a canonical model of mental representations, we can use signature limits to identify whether an individual's performance on a particular task involves efficient or flexible mindreading systems (1, 2).

AN ALTERNATIVE: THE EARLY MINDREADING ACCOUNT

The two-systems account contrasts with an approach suggesting that people have a unitary early developing (and possibly innate) psychological reasoning system that parses mental states from behavior. According to the early mindreading account, infants and young children succeed in violation-of-expectation or anticipatory-looking tasks because those tasks only require representing beliefs (10). Additional processes are involved in tasks that typically require responding verbally to a question; 3-year-olds also need to select between different responses to the test question and inhibit a default to answer from their own knowledge. The additional processes overwhelm 3-year-olds' limited executive functioning, masking their innate belief-reasoning competence.

According to the early mindreading account, some experiments suggest that 17- to 18-month-olds can attribute false beliefs about identity (29, 30). However, these experiments could just as well suggest that infants are tracking beliefs about the types of objects present rather than about numerical identity (2). That said, one study (31) suggested that infants' mindreading may be relatively sophisticated. Seventeen-month-olds watched a thief attempt to steal a preferred object (a rattling toy) when its owner was momentarily absent by replacing the toy with a less-preferred object (a toy that did not rattle). Infants looked longer when the thief replaced the preferred object with a silent toy that did not look like the original toy than when the thief replaced it with a silent toy that resembled the rattling toy. The authors postulated that infants ascribed to the thief an intention to implant in the owner a false belief about the identity of the substituted toy. They further suggested that infants made such ascriptions only when the replacement involved a toy that looked like the original toy and when the owner did not test whether the toy rattled when she returned to the room.

However, these explanations (31) also require that we believe that infants take the thief to be strikingly inept: Despite having the opportunity to simply pilfer from a closed box known to contain at least three rattling toys, the thief apparently engaged in an elaborate deception that would be uncovered whenever the replaced toy was shaken and that would easily identify the thief as the sole suspect. A further difficulty is that factors unrelated to the thief's mental states vary between conditions (e.g., the frequencies with which toys that visually match a toy present during the final phase of the test rattled). These considerations indicate that further evidence is needed to support the claim that people's early mindreading capacity enables them to ascribe intentions concerning false beliefs involving numerical identity.

In support of the early mindreading account, performance issues may also explain findings showing nonautomatic attribution of beliefs (32). With respect to the study mentioned earlier in which false beliefs were not ascribed automatically (11), researchers were concerned that the interval between the belief cues and belief questions was longer than the interval between the reality cues and reality questions. Adults might have been slower at attributing beliefs because they had to retrieve information about the agent's beliefs (which had been inferred automatically) from long-term memory when responding to unpredictable questions. Those researchers spotlighted another study (33), which argues that adults automatically inferred beliefs when the interval between belief cues and questions was shorter. However, in the other study, the context of the agent putting a marker on the wrong container just before the belief probe could just as well prompt adults to infer spontaneously (rather than automatically) the agent's false belief as a relevant explanation for her mistaken endorsement. Indeed, in a third study (34), task context motivated adults to make spontaneous inferences about an agent's beliefs and maintain them over time, even though they did not need to do so. However, in the absence of such motivation, participants did *not* automatically make belief inferences, even when the stimulus afforded such inferences.

The broader developmental evidence is also not entirely consistent with the explanation that contradictions in responses to false-belief scenarios reflect completely incidental demands on executive processing. Cultural differences in inhibitory control are not linked to corresponding differences in performance on standard false-belief tasks (35). Three-year-olds do not find selectionless false-belief tasks easier than standard false-belief tasks (36). It is also unclear why certain indirect tasks (e.g., the violation-of-expectation paradigm) are assumed to be free of inhibition demands when infants apparently face the same problem of controlling a default reading of the situation in terms of where the object is actually located to track beliefs instead (36). The notion of underlying competence in reasoning about beliefs being masked by incidental task demands to inhibit a tendency to answer from one's own knowledge would also need to be stretched to account for interference effects on reality judgments. In studies (37, 38), adults and children found it difficult to even hold others' false beliefs in mind, resulting in slower and incorrect judgments about reality. These considerations suggest that constraints on information processing play a deeper and more nuanced developmental role in the construction, maintenance, and use of belief concepts, in addition to lasting roles in the mature mindreading system (12). The two-systems account fits more optimally with the research that has looked at mindreading.

CONCLUSIONS

The two-systems approach to mindreading is motivated theoretically, and we are starting to see its predictions tested and confirmed. This approach is committed to an efficient system during infancy that is limited by the (minimal) model of mental representations that it relies on. In studies of different ages, populations, and paradigms, an efficient system tracking belief-like states can handle some visual-perspective and false-belief problems, but not others. Researchers need to map the terrain of the efficient (vs. flexible) mindreading system, and determine whether it is limited to handling certain kinds of agents, desire-like states, trait impressions, and perspective-based utterances of low complexity (1, 39, 40). Studying the temporal course of behavioral and neural activity associated with tracking belief-like states versus ascribing belief in real settings will also illuminate circumstances in which information might pass between systems, and delineate precise moments when mindreading inferences are constructed, stored, and used. Thinking about the cognitive architecture of human mindreading as involving many systems, models, and signature limits may be necessary to making sense of dissociations between different response classes as well as between nonautomatic and automatic processing.

REFERENCES

1. Apperly, I. A., & Butterfill, S. A. (2009). Do humans have two systems to track beliefs and belief-like states? *Psychological Review*, *116*, 953–970. doi:10.1037/a0016923
2. Butterfill, S. A., & Apperly, I. A. (2013). How to construct a minimal theory of mind. *Mind & Language*, *28*, 606–637. doi:10.1111/mila.12036
3. Wimmer, H., & Perner, J. (1983). Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children's understanding of deception. *Cognition*, *13*, 103–128. doi:10.1016/0010-0277(83)90004-5
4. 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. doi:10.1111/1467-8624.00304
5. Rakoczy, H., Bergfeld, D., Schwarz, I., & Fizke, E. (2015). Explicit theory of mind is even more unified than previously assumed: Belief ascription and understanding aspectuality emerge together in development. *Child Development*, *86*, 486–502. doi:10.1111/cdev.12311
6. Clements, W. A., & Perner, J. (1994). Implicit understanding of belief. *Cognitive Development*, *9*, 377–395. doi:10.1016/0885-2014(94)90012-4
7. Low, J. (2010). Preschoolers' implicit and explicit false-belief understanding: Relations with complex syntactical mastery. *Child Development*, *81*, 579–615. doi:10.1111/j.1467-8624.2009.01418.x
8. Ruffman, T., Garnham, W., Import, A., & Connolly, D. (2001). Does eye gaze indicate knowledge of false belief? Charting transitions in knowledge. *Journal of Experimental Child Psychology*, *80*, 201–224. doi:10.1006/jecp.2001.2633
9. Onishi, K. H., & Baillargeon, R. (2005). Do 15-month-old infants understand false beliefs? *Science*, *308*, 255–258. doi:10.1126/science.1107621

10. Baillargeon, R., Scott, R. M., & He, Z. (2010). False-belief understanding in infants. *Trends in Cognitive Sciences*, *14*, 110–117. doi:10.1016/j.tics.2009.12.006
11. Apperly, I. A., Riggs, K. J., Simpson, A., Chiavarino, C., & Samson, D. (2006). Is belief reasoning automatic? *Psychological Science*, *17*, 841–844. doi:10.1111/j.1467-9280.2006.01791.x
12. Apperly, I. A. (2010). *Mindreaders: The cognitive basis of "theory of mind."* Hove, UK: Psychology Press.
13. Schneider, D., Nott, Z. E., & Dux, P. E. (2014). Task instructions and implicit theory of mind. *Cognition*, *133*, 43–47. doi:10.1016/j.cognition.2014.05.016
14. Van der Wel, R. P. R. D., Sebanz, N., & Knoblich, G. (2014). Do people automatically track others' beliefs? Evidence from a continuous measure. *Cognition*, *130*, 128–133. doi:10.1016/j.cognition.2013.10.004
15. Carey, S. (2009). *The origins of concepts*. New York, NY: Oxford University Press.
16. Schneider, D., Lam, R., Bayliss, A. P., & Dux, P. E. (2012). Cognitive load disrupts theory of mind processing. *Psychological Science*, *23*, 842–847. doi:10.1177/0956797612439070
17. Thoermer, C., Sodian, B., Vuori, M., Perst, H., & Kristen, S. (2012). Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. *British Journal of Developmental Psychology*, *30*, 172–187. doi:10.1111/j.2044-835X.2011.02067.x
18. Samson, D., Apperly, I. A., Braithwaite, J. J., Andrews, B. J., & Bodley Scott, S. E. (2010). Seeing it their way: Evidence for rapid and involuntary computation of what other people see. *Journal of Experimental Psychology: Human Perception and Performance*, *36*, 1255–1266. doi:10.1037/a0018729
19. Furlanetto, T., Becchio, C., Samson, D., & Apperly, I. A. (2016). Altercentric interference in Level 1 visual perspective-taking reflects the ascription of mental states, not submentalizing. *Journal of Experimental Psychology: Human Perception and Performance*, *42*, 158–163. doi:10.1037/xhp0000138
20. Santiesteban, I., Catmur, C., Coughlan Hopkins, S., Bird, G., & Heyes, C. (2014). Avatars and arrows: Implicit mentalizing or domain-general processing? *Journal of Experimental Psychology: Human Perception and Performance*, *40*, 929–937. doi:10.1037/a0035175
21. 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
22. Surtees, A. D. R., Samson, D., & Apperly, I. A. (2016). Unintentional perspective-taking calculates whether something is seen, but not how it is seen. *Cognition*, *148*, 97–105. doi:10.1016/j.cognition.2015.12.010
23. Mozuraitis, M., Chambers, C. G., & Daneman, M. (2015). Privileged versus shared knowledge about object identity in real-time referential processing. *Cognition*, *142*, 148–165. doi:10.1016/j.cognition.2015.05.001
24. Low, J., & Watts, J. (2013). Attributing false beliefs about object identity reveals a signature blind spot in humans' efficient mindreading system. *Psychological Science*, *24*, 305–311. doi:10.1177/0956797612451469
25. Wang, B., Hadi, N. H., & Low, J. (2015). Limits on efficient human mindreading: Convergence across Chinese adults and Semai children. *British Journal of Psychology*, *106*, 724–740. doi:10.1111/bjop.12121
26. Fizke, E., Butterfill, S. A., & Rakoczy, H. (2013, April). *Toddlers' understanding of false belief about an object's identity*. Poster presented at the biennial meeting of the Society for Research in Child Development, Seattle, WA.
27. Cacchione, T., Schaub, S., & Rakoczy, H. (2013). Fourteen-month-old infants infer the continuous identity of objects on the basis of non-visible causal properties. *Developmental Psychology*, *49*, 1325–1329. doi:10.1037/a0029746
28. Buttelmann, D., Carpenter, M., & Tomasello, M. (2009). Eighteen-month-old infants show false belief understanding in an active helping paradigm. *Cognition*, *112*, 337–342. doi:10.1016/j.cognition.2009.05.006
29. Buttelmann, F., Suhrke, J., & Buttelmann, D. (2015). What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. *Journal of Experimental Child Psychology*, *131*, 94–103. doi:10.1016/j.jecp.2014.11.009
30. Scott, R. M., & Baillargeon, R. (2009). Which penguin is this? Attributing false beliefs about object identity at 18 months. *Child Development*, *80*, 1172–1196. doi:10.1111/j.1467-8624.2009.01324.x
31. Scott, R. M., Richman, J. C., & Baillargeon, R. (2015). Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning. *Cognitive Psychology*, *82*, 32–56. doi:10.1016/j.cogpsych.2015.08.003
32. Carruthers, P. (2015). Mindreading in adults: Evaluating two-systems views. *Synthese*, Advanced online publication. doi:10.1007/s11229-015-0792-3
33. Cohen, A., & German, T. (2009). Encoding of others' beliefs without overt instruction. *Cognition*, *111*, 356–363. doi:10.1016/j.cognition.2009.03.004
34. Back, E., & Apperly, I. A. (2010). Two sources of evidence on the non-automaticity of true- and false-belief ascription. *Cognition*, *115*, 54–70. doi:10.1016/j.cognition.2009.11.008
35. Sabbagh, M., Xu, F., Carlson, S., Moses, L., & Lee, K. (2006). The development of executive functioning and theory-of-mind: A comparison of Chinese and U.S. preschoolers. *Psychological Science*, *17*, 74–81. doi:10.1111/j.1467-9280.2005.01667.x
36. Wellman, H. M. (2014). *Making minds*. New York, NY: Oxford University Press.
37. Apperly, I. A., Back, E., Samson, D., & France, L. (2008). The cost of thinking about false beliefs: Evidence from adult performance on a non-inferential theory of mind task. *Cognition*, *106*, 1093–1108. doi:10.1016/j.cognition.2007.05.005
38. Low, J., & Simpson, S. (2012). Effects of labeling on preschoolers' explicit false-belief performance: Outcomes of cognitive flexibility or inhibitory control? *Child Development*, *83*, 1072–1084. doi:10.1111/j.1467-8624.2012.01738.x
39. Wang, J. J., Ali, M., Frisson, S., & Apperly, I. A. (2015). Language complexity modulates 8- and 10-year-olds' success at using their theory of mind abilities in a communication task. *Journal of Experimental Child Psychology*, Advanced online publication. doi:10.1016/j.jecp.2015.09.006
40. Schneider, D., & Low, J. (2016). Efficient versus flexible mentalizing in complex social settings: Exploring signature limits. *British Journal of Psychology*, *107*, 26–29. doi:10.1111/bjop.12165