Title
Early False-Belief Understanding

Journal
Trends in Cognitive Sciences, 21(4)

ISSN
1364-6613

Authors
Scott, Rose M
Baillargeon, Renée

Publication Date
2017-04-01

DOI
10.1016/j.tics.2017.01.012

Early False-belief Understanding

Rose M. Scott[a,*]

Renée Baillargeon[b]

[a]Psychological Sciences, University of California Merced, 5200 North Lake Road, Merced, California, 95343, United States of America

[b]Department of Psychology , University of Illinois, 603 E. Daniel St., Champaign, Illinois, 61820, United States of America

*Corresponding author. Email: rscott@ucmerced.edu

Keywords: social cognition; theory of mind; psychological reasoning; false-belief understanding

Abstract

Intense controversy surrounds the question of when children first understand that others can hold false beliefs. Results from traditional tasks suggest that false-belief understanding does not emerge until about 4 years of age and constitutes a major developmental milestone in social cognition. In contrast, results from nontraditional tasks, which have steadily accumulated over the past 10 years, suggest that false-belief understanding is already present in infants (under age 2) and toddlers (ages 2-3) and thus forms an integral part of social cognition from early in life. Here we first present an overview of the findings from nontraditional tasks. We then return to traditional tasks and argue that processing difficulties, rather than limitations in false-belief understanding, account for young children's failure at these tasks.

When does false-belief understanding emerge?

Adults routinely make sense of others' actions by inferring the mental states that underlie these actions; this ability is variously referred to as psychological reasoning, mindreading, or exhibiting a theory of mind. Much of the research on the development of this ability has focused on false-belief understanding, the capacity to understand that agents can be mistaken, or hold false beliefs, about the world. To demonstrate false-belief understanding, children must grasp, at least intuitively, the representational nature of the mind: They must realize that beliefs are internal representations rather than direct reflections of reality and, as such, can be inaccurate.

When are children first able to attribute false beliefs and other counterfactual mental states to others? This question has generated intense controversy because different false-belief tasks have suggested different answers, leading to different characterizations of the development of false-belief understanding and psychological reasoning more generally.

Traditionally, early false-belief understanding was assessed using *elicited-prediction* tasks, in which children are asked a test question that requires them to predict the behavior of an agent who holds a false belief [1-2]. In a well-known task [1], children heard a story enacted with props: Sally hid her marble in a basket and then left; in her absence, Anne moved the marble to a nearby box. Sally then returned, and children were asked, "Where will Sally look for her marble?" Most 4-year-olds correctly predicted that Sally would look for her marble in the basket; in contrast, most 3-year-olds incorrectly predicted that Sally would look in the box, as though unable to understand that Sally would hold a false belief about the marble's location. This developmental pattern—from below-chance to above-chance performance—was confirmed with other false-belief scenarios [3-4] and was widely replicated in cultures around the world [5-6]. Based on these results, many researchers concluded that a fundamental change takes place in psychological reasoning at about 4 years of age, when children begin to grasp the representational nature of the mind and become capable of understanding that agents can hold and act on false beliefs. False-belief understanding thus came to be viewed as a major developmental milestone that is not achieved until the preschool years and heralds a more advanced form of social cognition [2-8].

This conclusion was empirically challenged, however, by the discovery that 15-month-old infants demonstrated false-belief understanding when tested with a violation-of-expectation task [9] (violation-of-expectation tasks take advantage of infants' natural tendency to look longer at events that violate, as opposed to confirm, their expectations). Infants first saw an agent hide her toy in box-A as opposed to box-B. Next, infants received one of several belief-induction trials in which the agent came to hold either a true or false belief about the toy's location. In the subsequent test trial, the agent reached into either box-A or box-B and then paused. Infants expected the agent to reach into whichever box she believed contained the toy, regardless of whether her belief was true or false, and they detected a violation if she reached into the other box instead.

These results suggested that infants already attribute false beliefs to agents, calling into question the conclusion that false-belief understanding is not achieved until about 4 years of age. This finding launched a new wave of research as investigators from different laboratories devised a wide array of novel tasks to assess early false-belief understanding. We refer to these as nontraditional tasks, to distinguish them from the traditional, elicited-prediction tasks described above.

To date, over 30 published reports using nontraditional tasks have provided positive evidence of false-belief understanding in infants (under age 2) and toddlers (ages 2-3). In this article, we first present an overview of these findings and then discuss how they can be reconciled with the findings from traditional tasks.

Nontraditional false-belief tasks

Nontraditional tasks do not require answering a test question about the likely behavior of a mistaken agent: They use alternative ways of assessing children's understanding of the agent's false belief. Nontraditional tasks vary considerably in the paradigms they use, the false-belief scenarios they present, and the linguistic demands they impose. They have also been used, to similar effect, in different cultures.

Different paradigms

Nontraditional tasks can be divided into two categories: spontaneous-response and elicited-intervention tasks [10].

Spontaneous-response tasks. In spontaneous-response tasks, children watch a scene in which an agent comes to hold a false belief, and their understanding of this belief is inferred from their spontaneous responses to the unfolding scene. To date, evidence of early false-belief understanding has been obtained with seven types of spontaneous responses, some behavioral and others neural.

Focusing first on behavioral responses, when an agent holds a false belief about a scene, infants age 7 months and older have been found to look significantly longer when the agent acts in a manner that is inconsistent, as opposed to consistent, with this belief (*violation-of-expectation* tasks) [9, 11-24]. When an agent falsely believes that a desired object is in location-A, infants age 17 months and older visually anticipate that the agent will approach location-A (*anticipatory-looking* tasks) [25-32], and infants age 18 months and older spontaneously point to inform the agent that the object has been moved to another location or has been replaced with an aversive object (*anticipatory-pointing* tasks) [33-34]. When told a false-belief story accompanied by pictures, toddlers age 2.5 years and older look preferentially at the final picture that correctly, as opposed to incorrectly, completes the story (*preferential-looking* tasks) [35]. Toddlers age 2.5 years and older also express more tension in their facial expressions when an agent who is approaching a container is mistaken, as opposed to ignorant, about its contents (*affective-response* tasks) [36].

Turning next to neural responses, when an agent falsely believes that a container holds a desired object, 6-month-olds expect the agent to search for the object: Electroencephalographic (EEG) measurement shows an increase in sensorimotor alpha-band suppression (a neural correlate of action prediction), which is absent when the agent falsely believes that the container is empty (*neural action-prediction* task) [37]. Finally, when an agent falsely believes that an object is behind an occluder (unbeknownst to the agent, the object disintegrates once occluded), 8-month-olds encode the agent's false belief about the continued presence of the object: EEG measurement shows an increase in temporal gamma-band activation (a neural correlate of sustained object representation during occlusion), which is absent when the agent witnesses the object's disintegration (*neural sustained-representation* task) [38].

Elicited-intervention tasks. In elicited-intervention tasks, children watch a scene in which an agent comes to hold a false belief, and then they are prompted to perform some action for the agent; to succeed, children must take into account the agent's false belief [39-43]. For instance, in one task [39], an experimenter first showed 18-month-olds how to lock and unlock box-A and box-B; the boxes were left unlocked. Next, an agent entered the room, hid his toy in box-A, and then left. While he was gone, the experimenter moved the toy to box-B and locked both boxes. When the agent returned, he tried vainly to open box-A. When prompted to help him

("Go on, help him!"), most infants approached box-B, suggesting that they understood he wanted his toy and falsely believed it was still in box-A. In another task [43], 17-month-olds watched as an agent hid two distinct toys in box-A and box-B and then left. In her absence, an experimenter switched the toys' locations. When the agent returned, she pointed to box-A, said she wanted the toy in it, and asked infants, "Can you get it for me?" Most infants approached box-B, suggesting that they realized the agent held a false belief about which toy was in which box and that they understood which toy she wanted.

Different false-belief scenarios

Nontraditional tasks have employed a wide range of false-belief scenarios that differ along many dimensions.

Types of false beliefs. Nontraditional tasks have produced evidence that infants and toddlers can attribute a wide variety of false beliefs to agents, including false beliefs about the presence [14, 37], location [9, 39], identity [17, 41], and obvious and non-obvious properties [18, 36] of objects, as well as false beliefs about the moral characters of other agents [12].

Nontraditional tasks also differ in the causal processes that give rise to agents' false beliefs. In some cases, agents gain information when witnessing an event, but this information becomes outdated in their absence. For example, when an agent saw a desired object move into location-A, and in the agent's absence the object moved to location-B or left the scene, infants age 6 months and older expected the returning agent to falsely believe the object was still in location-A [9, 37]. In other cases, agents are led by a series of similar events to assume, wrongly, that the next event will follow the same pattern. For example, after an agent saw a desired object being hidden in location-A four times, 13-month-olds expected the agent to falsely assume the object had again been hidden in location-A [22]. Similarly, after an agent was shown that three boxes contained a block, 18-month-olds expected the agent to falsely assume a fourth box also contained a block [40]. In yet other cases, agents are led by a general expectation about the world to draw an inference that happens not to hold true in the scene. For example, after seeing an experimenter shake object-A to produce a rattling sound, 18-month-olds expected an agent who wanted to produce the same effect to mistakenly select a similar object-A over a dissimilar object-B, on the general expectation that similar objects are more likely to share non-obvious properties [18]. Likewise, when an agent announced that she wanted to color, 2.5-year-olds expected her to reach for a box of crayons as opposed to a box of cheerios, on the general expectation that commercial packages usually hold their depicted contents [13].

Finally, in some nontraditional tasks, the agent holds not just one false belief, but two causally interlocking false beliefs. For example, when two toys were placed in location-A and location-B, and perceptual or contextual cues misled the agent about the identity of the toy in location-A, infants age 14 months and older expected the agent to falsely infer that the other toy was in location-B [17, 20].

Belief-based behaviors. Nontraditional tasks have produced evidence that infants and toddlers can reason about a wide range of behaviors by agents with false beliefs. These behaviors include: (a) physical actions, such as where a mistaken agent will search for a desired object or which object she will select to produce a desired effect [9, 18]; (b) social interactions, such as whether a mistaken agent will continue to interact positively with another agent [12]; (c) verbal statements, such as which object a mistaken agent intends to label or request [42-43]; and (d) emotional responses, such as how an agent will react upon discovering she was mistaken [16]. In this last task, 20-month-olds first watched an agent play with two rattling toys. In her absence, an experimenter manipulated one of the toys to render it silent. When the agent returned and shook the silent toy, infants expected her to look surprised and looked significantly longer if she looked satisfied instead (this pattern reversed if the agent knew one of the toys had been manipulated).

Infants can reason not only about the behaviors of agents who hold false beliefs, but also about the behaviors of deceptive agents who seek to implant false beliefs. In a task with 17-month-olds [19], a thief attempted to secretly steal a desirable rattling toy during its owner's absence by substituting a less desirable silent toy (Figure 1). Results indicated that infants understood that the thief (a) sought to implant in the owner a false belief about the identity of the silent toy and (b) could achieve this deceptive goal only by substituting a silent toy that was visually identical to the rattling toy (otherwise the owner would detect the substitution as soon as she saw the silent toy). Additional results indicated that infants expected the returning owner to be deceived by this substitution and to store the silent toy in her treasure box alongside her other rattling toys.

Different linguistic demands

While nontraditional tasks with infants are typically non-verbal, those with toddlers can be either non-verbal or verbal. Some verbal nontraditional tasks make linguistic demands comparable to those of traditional tasks. In one task [35], for example, 2.5-year-old toddlers heard a story about a character named Emily who hid an apple in one of two locations; in her absence, the apple was moved to the other location. The story was accompanied by a large picture book; each double-page showed one picture that matched the story and one that did not. The story ended with the line, "Emily is looking for her apple", and in the final double-page one picture showed Emily searching for her apple where she falsely believed it to be (*original-location* picture) and the other picture showed her searching for her apple in its current location (*current-location* picture). Results indicated that (a) as the story unfolded, toddlers looked preferentially at the matching picture on each double-page, suggesting that they had no difficulty following the story, and (b) on the final double-page, toddlers looked preferentially at the original-location picture, suggesting that they understood that Emily would falsely believe her apple was still in its original location.

Other verbal nontraditional tasks incorporate test questions similar to those asked in traditional tasks (e.g., "Where will Sally look for her marble?"). In some tasks, instead of directing the test question at the child, the experimenter utters it in a self-addressed manner, as though thinking out loud, and investigators measure whether toddlers visually anticipate which location Sally will search [25-26]. In other tasks, the test question is directed at an adult "bystander," and researchers measure whether toddlers look significantly longer when the adult answers incorrectly and points to the toy's current as opposed to original location [35]. Both types of tasks have yielded positive results with toddlers age 2.5 years and older. Thus, although toddlers fail to demonstrate false-belief understanding when asked the test question directly (as shown in traditional tasks), they succeed when they merely overhear this question.

Different cultures

Nontraditional tasks have produced evidence of early false-belief understanding not only in Western cultures, as reviewed above, but also in non-Western cultures [11]. Positive results were obtained with three spontaneous-response tasks—a non-verbal violation-of-expectation task and verbal anticipatory-looking and preferential-looking tasks—in traditional non-Western communities: a Salar community in western China, a Shuar/Colono community in Ecuador, and a Yasawan community in Fiji. In each task, children performed similarly to children from the United States [18, 26, 35], suggesting that the capacity to attribute false beliefs emerges universally early in development.

Why do young children fail at traditional false-belief tasks?

How can we reconcile the strikingly divergent findings from traditional and nontraditional false-belief tasks? Why do children fail at traditional tasks until about 4 years of age, but succeed at nontraditional tasks beginning in the first year of life? There currently exist two broad views on this question.

According to the *fundamental-change* view, traditional tasks tap genuine false-belief understanding, which develops during the preschool years as a result of significant conceptual, executive-function, linguistic, and/or meta-representational advances [2-8, 44-51]. According to Perner and Roessler [48], for example, correctly answering test questions such as, "Where will Sally look for her marble?" "requires an intentional switch of perspectives not possible before 4 years of age" (p. 519). For some proponents of this view, the evidence from nontraditional tasks is open to low-level alternative interpretations that implicate no false-belief understanding [46, 49]; for other proponents, this evidence reveals only a minimal form of false-belief understanding [44-45, 47, 50-51] (Box 1). Either way, a major shift is thought to occur in the preschool years, which makes possible correct responses in traditional tasks.

According to the *substantial-continuity* view, traditional and nontraditional tasks tap the same genuine false-belief understanding, but traditional tasks are subject to greater processing difficulties [52-56]. Although proponents of this view differ in their descriptions of these difficulties (Box 2), they agree that false-belief reasoning emerges early in life and gradually becomes more efficient and more nuanced with age and experience.

Which of these views is correct? From our perspective, the extensive evidence reviewed in the previous section casts doubt on the fundamental-change view. As convergent findings of early false-belief understanding have steadily accumulated over the past 10 years, it has become increasingly unlikely that low-level alternative interpretations could account for all of these findings. Claims that infants and toddlers are limited to a minimal form of false-belief understanding are also difficult to accept. As our review of nontraditional tasks makes clear, early false-belief understanding is remarkably sophisticated. Infants and toddlers correctly reason about many different false-belief scenarios, and they express this understanding in behaviors ranging from looking and affective responses to spontaneous and elicited actions. Toddlers also correctly interpret test questions such as, "Where will Sally look for her marble?" when they merely overhear such questions. Given these various findings, it seems unlikely that traditional tasks tap an advanced form of false-belief understanding fundamentally distinct from that available to infants and toddlers.

Additional evidence against the fundamental-change view comes from experiments that tested one key prediction from the substantial-continuity view: If young children fail at traditional tasks due to processing difficulties, then reducing these difficulties should result in success before age 4. At least two sets of findings have provided support for this prediction.

Reducing inhibitory-control demands. Many attempts at reducing processing difficulties in traditional tasks have focused on inhibitory-control demands for two reasons. First, Leslie and his colleagues [57-58] proposed a computational model in which inhibitory control is necessary for children to express their false-belief understanding. When children are asked the test question, "Where will Sally look for her marble?", an incorrect prepotent response based on the marble's actual location is triggered (why this is so is currently debated, as discussed in Box 2); this prepotent response must then be inhibited for children to answer correctly based on Sally's false belief. Because young children's inhibitory control is immature [59], however, they cannot effectively suppress this response and thus mistakenly point to the marble's current location.

Second, correlational studies with preschoolers found a significant association between performance in traditional false-belief tasks and performance in tasks that measure conflict inhibitory control, the ability to suppress a prepotent response while activating a conflicting response (e.g., saying “day” when shown a picture of the moon) [60-62].

In line with this research, investigators have found that when inhibitory-control demands in traditional tasks are reduced by various means, 3.5- to 4-year-olds often succeed [63-69]. In one low-inhibition task, for example, children were asked where Sally would look *first* for her marble, underscoring the need to respond based on Sally’s belief rather than on the marble’s current location [67-69]. In another task, children helped the experimenter deceive Sally by moving her marble to a new location, again underscoring the importance of Sally’s false belief [64]. In yet another task, children were told that Sally believed her marble was in the basket when in fact both containers were empty or held other objects; because children did not know the marble’s actual location, the incorrect prepotent response triggered by the test question was weaker and easier to suppress [63].

Reducing total processing demands. In contrast to 3.5- to 4-year-olds, children age 3 and younger typically perform only at chance in low-inhibition tasks [8, 68-70]. In one task [70], for example, 2.5-year-old toddlers heard a story accompanied by a picture book: Emma found an apple in one of two containers, moved it to the other container, and then went outside to play with her ball; in her absence, her brother Ethan found the apple and took it away. Emma then returned to look for her apple. In the test trial, children were shown pictures of the two containers and were asked the test question, “Where will Emma look for her apple?” Because toddlers did not know the apple’s actual location, the weak prepotent response triggered by the test question should have been easier to suppress. Nevertheless, toddlers performed only at chance. Such negative results are often taken to support the fundamental-change view [8, 62]: Only transitional children who are approaching their fourth birthday and are well on their way to acquiring the genuine false-belief understanding necessary for success at traditional tasks can fully benefit from a reduction in inhibitory-control demands. However, alternative interpretations of these negative results consistent with the substantial-continuity view are also possible.

In line with their expanded-processing-demands account (discussed in Box 2), Setoh, Scott, and Baillargeon [70] suggested that children age 3 and younger might perform at chance in low-inhibition tasks simply because the total amount of concurrent processing demands in the tasks, though admittedly reduced, is still large enough to overwhelm young children's limited information-processing resources. In the low-inhibition task described above, for example, toddlers had to manage at least three processes: *false-belief representation* – as the story unfolded, children had to form and maintain a representation of Emma's false belief; *response generation* – when asked the test question, children had to interpret it, hold it in mind, and generate a response; and *inhibitory control* – children had to inhibit the weak incorrect prepotent response triggered by the test question to answer correctly. This analysis led to the following prediction: If toddlers performed at chance in the task because their limited information-processing resources were overwhelmed by the total amount of processing demands in the task, then they might succeed if this amount were further lowered by also reducing response-generation demands.

To test this prediction, two practice trials were interspersed among the story trials (Figure 2). In one practice trial, children saw an apple and a banana and were asked, "Where is Emma's apple?"; in the other, they saw a ball and a frisbee and were asked, "Where is Emma's ball?" In each case, toddlers were required to point to the matching picture. These trials thus gave children practice interpreting a “where” question and producing a response by pointing to one of two pictures. As predicted, toddlers now performed *above chance* in the test trial, pointing to the container Emma falsely believed held her apple. Additional experiments indicated that toddlers performed *at chance* if they received only one practice trial or if the practice trials differed in form from the test trial, rendering them less effective at reducing response-generation demands. Finally, toddlers performed *below chance* if the practice trials were embedded in a high-inhibition false-belief story, thereby increasing inhibitory-control demands.

Thus, across experiments, slight changes in the task’s processing demands led 2.5-year-old toddlers to perform above chance, at chance, or below chance, providing evidence for the claim that early failures at traditional tasks stem from processing difficulties. Together with the many positive results from nontraditional tasks with infants and toddlers reviewed in the previous section, these findings provide strong support for the substantial-continuity view.

Concluding Remarks

The research reviewed in this article suggests three conclusions. First, false-belief understanding emerges early in life and is robust and sophisticated, allowing young children to reason about a wide variety of false-belief scenarios and to express this understanding with a wide range of spontaneous and elicited responses. Second, young children fail at traditional tasks not because their false-belief understanding is limited but because (a) they lack sufficient skill at one or more of the extraneous processes in the tasks or (b) they lack sufficient information-processing resources to handle the tasks’ total processing demands. Third, reconciling findings from different false-belief tasks requires considering the full range of processes associated with each task, and this is true for both traditional and nontraditional tasks: Just as reducing processing difficulties in a traditional task can lead to success at younger ages [63-70], increasing processing difficulties in a nontraditional task can lead to success at later ages [71; see also 72].

These conclusions are consistent with prior evidence that multiple factors are related to success in traditional tasks, including (a) developmental changes in inhibitory control, working memory, and verbal ability [60-62, 73], and (b) practice at conversing and answering questions about mental states with parents or siblings [74-76]. These conclusions are also consistent with findings that young children with advanced inhibitory-control abilities, such as crib bilinguals [77] and Chinese preschoolers [78], do not perform above chance in traditional tasks. Inhibitory-control demands are not the only demands in these tasks, so reducing only these demands may not be sufficient to allow success [see also 79].

In sum, the evidence reviewed in this article indicates that infants' and toddlers' psychological-reasoning system allows them to represent counterfactual states as well as motivational and epistemic states. Contrary to what was traditionally thought, counterfactual-state reasoning does not constitute a major milestone in the development of psychological reasoning; it emerges early in life and, from the start, may contribute in several distinct ways to everyday social cognition (Box 3).

Acknowledgements

We thank Francesco Antilici, Dan Hyde, and three anonymous reviewers for helpful comments and discussions.

References

1. Baron-Cohen, S. et al. (1985) Does the autistic child have a “theory of mind”? Cognition 21, 37-46
2. Wimmer, H. and Perner, J. (1983) Beliefs about beliefs: Representation and constraining function of wrong beliefs in young children’s understanding of deception. Cognition 13, 103–128
3. Gopnik, A. and Astington, J.W. (1988) Children’s understanding of representational change and its relation to the understanding of false belief and the appearance-reality distinction. Child Dev. 59, 26-37
4. Perner, J. et al. (1987) Three-year-olds’ difficulty with false belief: The case for a conceptual deficit. Br. J. Dev. Psychol. 5, 125–137
5. Callaghan, T. et al. (2005) Synchrony in the onset of mental-state reasoning. Psychol. Sci. 16, 378–384
6. Wellman, H. et al. (2006) Scaling of theory-of-mind understandings in Chinese children. Psychol. Sci. 17, 1075-1081
7. Liu, D. et al. (2008) Theory of mind development in Chinese children: A meta-analysis of false-belief understanding across cultures and languages. Dev. Psychol. 44, 523–531
8. Wellman, H.M. et al. (2001) Meta-analysis of theory-of-mind development: The truth about false belief. Child Dev. 72, 655–684
9. Onishi, K.H. and Baillargeon, R. (2005) Do 15-month-old infants understand false beliefs? Science 308, 255–258

10. Baillargeon, R. et al. (2016) Psychological reasoning in infancy. Annu. Rev. Psychol. 67, 159-186

11. Barrett, H.C. et al. (2013) Early false-belief understanding in traditional non-Western Societies. Proc. R. Soc. B 280, 20122654

12. Choi, Y. and Luo, Y. (2015) 13-month-olds' understanding of social interactions. Psychol. Sci. 26, 274-283

13. He, Z. et al. (2011) False-belief understanding in 2.5-year-olds: Evidence from violation-of-expectation change-of-location and unexpected-contents tasks. Dev. Sci. 14, 292-305

14. Kovács, Á.M. et al. (2010) The social sense: Susceptibility to others' beliefs in human infants and adults. Science 330, 1830-1834

15. Luo, Y. (2011) Do 10-month-old infants understand others' false beliefs? Cognition 121, 289-298

16. Scott, R.M. (2017) Surprise! 20-month-old infants understand the emotional consequences of false beliefs. Cognition 159, 33-47

17. Scott, R.M. and Baillargeon, R. (2009) Which penguin is this? Attributing false beliefs about identity at 18 months. Child Dev. 80, 1172–1196

18. Scott, R.M. et al. (2010) Attributing false beliefs about non-obvious properties at 18 months. Cogn. Psychol. 61, 366-395

19. Scott, R.M. et al. (2015) Infants understand deceptive intentions to implant false beliefs about identity: New evidence for early mentalistic reasoning. Cogn. Psychol. 82, 32-56

20. Song, H. and Baillargeon, R. (2008) Infants' reasoning about others' false perceptions. Dev. Psychol. 44, 1789–1795

21. Song, H. et al. (2008) Can an agent's false belief be corrected through an appropriate communication? Psychological reasoning in 18-month-old infants. Cognition 109, 295–315

22. Surian, L. et al. (2007) Attribution of beliefs by 13-month-old infants. Psychol. Sci. 18, 580–586

23. Träuble, B. et al. (2010) Early theory of mind competencies: Do infants understand others' beliefs? Infancy 15, 434-444

24. Yott, J. and Poulin-Dubois, D. (2012) Breaking the rules: Do infants have a true understanding of false belief? Br. J. Dev. Psychol. 30, 156-171

25. Garnham, W.A. and Ruffman, T. (2001) Doesn't see, doesn't know: Is anticipatory looking really related to understanding of belief? Dev. Sci. 4, 94-100

26. He, Z. et al. (2012) 2.5-year-olds succeed at a verbal anticipatory-looking false-belief task. Br. J. Dev. Psychol. 30, 14-29

27. Meristo, M. et al. (2012) Belief attribution in deaf and hearing infants. Dev. Sci. 15, 633-640

28. Senju, A. et al. (2011) Do 18-month-olds really attribute mental states to others? A critical test. Psychol. Sci. 22, 878-880

29. Sodian, B. et al. (2016) Understanding of goals, beliefs, and desires predicts morally relevant theory of mind: A longitudinal investigation. Child Dev. 87, 1221-1232

30. Southgate, V. et al. (2007) Action anticipation through attribution of false belief by two-year-olds. Psychol. Sci. 18, 587–592

31. Surian, L. and Geraci, A. (2012) Where will the triangle look for it? Attributing false beliefs to geometric shapes at 17 months. Br. J. Dev. Psychol. 30, 30-44

32. Thoermer C. et al. (2012) Continuity from an implicit to an explicit understanding of false belief from infancy to preschool age. Br. J. Dev. Psychol. 30, 172-87

33. Knudsen, B. and Liszkowski, U. (2012) Eighteen- and 24-month-old infants correct others in anticipation of action mistakes. Dev. Sci. 15, 113-122

34. Knudsen, B. and Liszkowski, U. (2012) 18-month-olds predict specific action mistakes through attribution of false belief, not ignorance, and intervene accordingly. Infancy 17, 672-691

35. Scott, R.M. et al. (2012) False-belief understanding in 2.5-year-olds: Evidence from two novel verbal spontaneous-response tasks. Dev. Sci. 15, 181-193

36. Moll, H. et al. (2017) 2.5-year-olds express suspense when others approach reality with false expectations. Child Dev. 88, 114-122

37. Southgate, V. and Vernetti, A. (2014) Belief-based action prediction in preverbal infants. Cognition 130, 1-10

38. Kampis, D. et al. (2015) Neural signatures for sustaining object representations attributed to others in preverbal human infants. Proc. R. Soc. B. 282, 20151683

39. Buttelmann, D. et al. (2009) Eighteen-month-old infants show false belief understanding in an active helping paradigm. Cognition 112, 337–342

40. Buttelmann, D. et al. (2014) Eighteen-month-olds understand false beliefs in an unexpected-contents task. J. Exp. Child Psychol. 119, 120-126

41. Buttelmann, F. et al. (2015) What you get is what you believe: Eighteen-month-olds demonstrate belief understanding in an unexpected-identity task. J. Exp. Child Psychol. 131, 94-103

42. Carpenter, M. et al. (2002) A new false-belief test for 36-month-olds. Br. J. Dev. Psychol. 20, 393-420
43. Southgate, V. et al. (2010) Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. Dev. Sci. 16, 907-912
44. Butterfill, S. and Apperly, I. (2013) How to construct a minimal theory of mind. Mind Lang. 28, 606-637
45. De Bruin, L.C. and Newen, A. (2012) An association account of false belief understanding. Cognition 123, 240-259
46. Heyes, C. (2014) False belief in infancy: A fresh look. Dev. Sci. 17, 647-659
47. Low, J. et al. (2016) Cognitive architecture of belief reasoning in children and adults: A primer on the two-systems account. Child Dev. Perspect. 10, 184-189
48. Perner, J. and Roessler J. (2012) From infants' to children's appreciation of beliefs. Trends Cogn. Sci. 16, 519-525
49. Ruffman, T. (2014) To belief or not belief: children's theory of mind. Dev. Rev. 34, 265-293
50. Sodian, B. (2011) Theory of mind in infancy. Child Dev. Perspect. 5, 39-43
51. Wellman, H.M. (2014) *Making minds: How theory of mind develops*, Oxford University Press
52. Baillargeon, R. et al. (2010) False-belief understanding in infants. Trends Cogn. Sci. 14, 110-118
53. Carruthers, P. (2013) Mindreading in infancy. Mind Lang. 28, 141-172
54. Helming, K.A. et al. (2014) Making sense of early false-belief understanding. Trends Cogn. Sci. 18, 167-170

55. Baker, S.T. et al. (2016) Bayesian change-point analysis reveals developmental change in a classic theory of mind task. Cogn. Psychol. 91, 124-149

56. Scott, R.M. (in press) The developmental origins of false-belief understanding. Curr. Dir. Psychol. Sci.

57. Leslie, A.M. and Polizzi, P. (1998) Inhibitory processing in the false belief task: Two conjectures. Dev. Sci. 1, 247-53

58. Leslie, A.M. et al. (2004) Core mechanisms in 'theory of mind.' Trends Cogn. Sci. 8, 528-533

59. Diamond, A. (2013) Executive functions. Annu. Rev. Psychol. 64, 135-168

60. Carlson, S.M. and Moses, L.J. (2001) Individual differences in inhibitory control and children's theory of mind. Child Dev. 72, 1032-1053

61. Hala, S. et al. (2003) Executive function and false-belief understanding in preschool children: Two tasks are harder than one. J. Cogn. Dev. 4, 275-298

62. Devine, R.T. and Hughes, C. (2014) Relations between false belief understanding and executive function in early childhood: A meta-analysis. Child Dev. 85, 1777-1794

63. Bartsch, K. (1996) Between desires and beliefs: Young children's action predictions. Child Dev. 67, 1671-1685

64. Hala, S. and Chandler, M. (1996) The role of strategic planning in accessing false-belief understanding. Child Dev. 67, 2948-2966

65. Kikuno, H. et al. (2007) How do young children process beliefs about beliefs? Evidence from response latency. Mind Lang. 22, 297-316

66. Mitchell, P. and Lacohée, H. (1991) Children's early understanding of false belief. Cognition 39, 107-127

67. Siegal, M. and Beattie, K. (1991) Where to look first for children's knowledge of false beliefs. Cognition 38, 1-12

68. Surian, L. and Leslie, A.M. (1999) Competence and performance in false belief understanding: a comparison of autistic and normal 3-yr-old children. Br. J. Dev. Psychol. 17, 141-155

69. Yazdi, A.A. et al. (2006) Competence and performance in belief-desire reasoning across two cultures: the truth, the whole truth, and nothing but the truth about false belief. Cognition 100, 343-368

70. Setoh, P. Scott, R.M. et al. (2016) Two-and-a-half-year-olds succeed at a traditional false-belief task with reduced processing demands. Proc. Natl. Acad. Sci. 113, 13360-13365

71. Scott, R.M. and Roby, E. (2015) Processing demands impact 3-year-olds' performance in a spontaneous-response task: New evidence for the processing-load account of early false-belief understanding. PLoS ONE 10, e0142405

72. Schneider, D. et al. (2012) Cognitive load disrupts implicit theory-of-mind processing. Psychol. Sci. 23, 842-847

73. Milligan, K. et al. (2007) Language and theory of mind: Meta-analysis of the relation between language ability and false-belief understanding. Child Dev. 78, 622-646

74. Brown, J.R. et al. (1996) Why talk about mental states? The significance of children's conversations with friends, siblings, and mothers. Child Dev. 67, 836-849

75. Ensor, R., and Hughes, C. (2008). Content or connectedness? Mother-child talk and early social understanding. Child Dev. 79, 201-216

76. Ruffman, T. et al. (2002) The relation between children's and mothers' mental state language and theory-of-mind understanding. Child Dev. 73, 734-751

77. Kovács, Á.M. (2009) Early bilingualism enhances mechanisms of false-belief reasoning. Dev. Sci. 12, 48-54

78. Sabbagh, M. et al. (2006) The development of executive functioning and theory of mind. Psychol. Sci. 17, 74-81

79. Duh, S. et al. (2016) Theory of mind and executive function in Chinese preschool children. Dev. Psychol. 52, 582-591

80. Ruffman, T. and Perner, J. (2005) Do infants really understand false belief? Response to Leslie. Trends Cogn. Sci. 9, 462-463

81. Perner, J. (2010) Who took the cog out of cognitive science? Mentalism in an era of anti-cognitivism. In *Cognition and neuropsychology: International perspectives on psychological science* (Frensch, P.A. and Schwarzer, R. eds), pp. 241-261, Psychology Press

82. Scott, R.M. and Baillargeon, R. (2014) How fresh a look? A reply to Heyes. Dev. Sci. 17, 660-664

83. Scott, R.M. (2014) Post hoc versus predictive accounts of children's theory of mind: a reply to Ruffman. Dev. Psychol. 3, 300-304

84. Baillargeon, R. et al. (2015) Psychological and sociomoral reasoning in infancy. In *APA Handbook of Personality and Social Psychology: Vol. 1. Attitudes and Social Cognition* (Mikulincer, M. and Shaver, P.R., eds, Borgida, E. and Bargh, J.A., assoc. eds), pp 79-150, American Psychological Association

85. Hyde, D.C. and Spelke, E.S. (2011) Neural signatures of number processing in human infants: evidence for two core systems underlying non-verbal numerical cognition. Dev. Sci. 14, 360-371

86. Xu, Y. and Chun, M.M. (2009) Selecting and perceiving multiple visual objects. Trends Cogn. Sci. 13, 167-174

87. Bardi, L. et al. (in press) Brain activation for spontaneous and explicit false belief tasks overlaps: New fMRI evidence on belief processing and violation of expectation. Soc. Cogn. Affect. Neurosci. DOI: 10.1093/scan/nsw143

88. Hyde, D.C. et al. (2015) Human temporal-parietal junction spontaneously tracks others' beliefs: A functional near-infrared spectroscopy study. Hum. Brain Mapp. 36, 4831-4846

89. Kovács, Á.M. et al. (2014) Are all beliefs equal? Implicit belief attributions recruiting core brain regions of theory of mind. PLOS One 9, e106558

90. Carruthers, P. (2016) Two systems for mindreading? Rev. Philos. Psychol. 7,141-162

91. Wang, L. and Leslie, A.M. (2016) Is implicit theory of mind the "real deal"? The own-belief/true-belief default in adults and young preschoolers. Mind Lang. 31, 147-176

92. Helming, K. et al. (2016) Solving the puzzle about early belief-ascription. Mind Lang. 31, 438-469

93. Rubio-Fernández, P. and Geurts, B. (2013) How to pass the false-belief task before your fourth birthday. Psychol. Sci. 24, 27-33

94. Rubio-Fernández, P. and Geurts, B. (2016) Don't mention the marble! The role of attentional processes in false-belief tasks. Rev. Philos. Psychol. 7, 835-850

95. Westra, E. and Carruthers, P. (2017) Pragmatic development explains the theory-of-mind scale. Cognition 158, 165-176

96. Roby, E.R. and Scott, R.M. (2016) Rethinking the relationship between social experience and false-belief understanding: A mentalistic account. Front. Psychol. 7, 1721

97. Graham, J. et al. (2013) Moral foundations theory: The pragmatic validity of moral pluralism. Adv. Exp. Soc. Psychol. 47, 55-130

98. Rai, T.S. and Fisk, A.P. (2011) Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. Psychol. Rev. 118, 57-75

99. Baillargeon, R. et al. (2013) False-belief understanding and why it matters: The social-acting hypothesis. In *Navigating the social world: What infants, children, and other species can teach us* (eds. Banaji, M.R. and Gelman, S.A.), pp. 88-95, Oxford University Press

100. Krupenye, C. et al. (2016) Great apes anticipate that other individuals will act according to false beliefs. Science 354, 110-114

Box 1: Fundamental-change accounts

Proponents of the fundamental-change view have offered two types of accounts for the findings of nontraditional tasks.

Nonmentalistic accounts. Nonmentalistic accounts argue that infants cannot represent mental states and that their responses in nontraditional tasks stem from low-level processes. In some accounts, infants' responses reflect perceptual novelty [46, 80]. In initial trials, infants encode "configurations of persons relating to objects" [80, p. 462] or, more primitively, configurations of "colours, shapes, and movements" [46, p. 648]; in subsequent trials, infants look longer when novel configurations deviate from previous encodings. Other accounts argue that infants bring to the laboratory behavioral rules [49, 81] that "capture the workings of the mind without mentioning the mind" [81, p. 259]. In everyday life, infants detect mind-blind statistical rules for how agents behave in particular situations. When infants encounter similar situational conditions in nontraditional tasks, they retrieve the appropriate rules and look longer when observed actions deviate from predicted actions.

There are two reasons to doubt nonmentalistic accounts. First, these accounts are inconsistent with the wealth of evidence, accumulated over the past 20 years, that infants engage in psychological reasoning [10, 82-84]. Second, many nontraditional tasks include control conditions that challenge low-level interpretations, because across conditions infants respond differently to similar test events [17-19, 28]. The study depicted in Figure 1, for example, included a control condition in which the owner always shook her toy when she returned in the familiarization trials, before storing or discarding it. Infants now looked equally during the matching and non-matching test trials, as neither substitution could deceive the owner.

Two-system accounts. Other accounts grant infants a minimal capacity for reasoning about mental states, but assume that the *early-developing* system that makes possible success at nontraditional tasks is distinct from, and considerably more primitive than, the *late-developing* system that emerges around age 4 and enables success at traditional tasks [44-45, 47, 50-51]. In Butterfill and Apperly's two-system account [44], for example, the early-developing system cannot represent false beliefs per se: instead, it tracks belief-like "registrations." Upon encountering an object, an agent registers its location and properties; by tracking this registration—even if it becomes outdated in the agent's absence—the early-developing system can predict the agent's actions (e.g., an agent will search for her toy where she last registered it).

There are several reasons to doubt that two systems with distinct neurological substrates and computational capacities underlie success in traditional and nontraditional tasks. In the case of numerical cognition, for example, there is overwhelming evidence that the object-tracking and analog-number systems activate distinct brain regions and perform distinct computations [85-86]. The case of false-belief understanding is very different, however. First, neuroscientific investigations with adults indicate that traditional and nontraditional tasks engage anatomically similar regions within the temporal-parietal junction [87-89]. Second, claims about the early-developing system's signature limits (e.g., an inability to track false beliefs about identity) have been overturned [17, 19-20, 41; see also 90]. Finally, as discussed in the text, even toddlers succeed at traditional tasks when processing demands are sufficiently reduced [70].

Box 2: Substantial-continuity accounts

Proponents of the substantial-continuity view differ in their accounts of which processing difficulties lead young children to fail at traditional tasks such as the Sally-Anne task.

Inhibitory-demands account. In this account, children fail to attribute a false belief to Sally due to limited inhibitory control [58, 91]. When children are asked the test question, their psychological-reasoning system suggests two possible beliefs for attribution, one corresponding to Sally's false belief and one corresponding to children's own true belief about the marble's location. A selection process then reviews these beliefs and attributes the second one by default (as agents' beliefs are usually true), unless it has sufficient inhibitory power suppress this *true-belief* bias.

Pragmatic accounts. In pragmatic accounts, children correctly attribute a false belief to Sally but misinterpret the test question due to limited pragmatic skills [54, 67, 69, 92-95]. In some of these accounts, the experimenter's mention of the marble leads children to focus on its current location or on the knowledge they share with the experimenter about this location (*referential* bias) [54, 92-94]. In other accounts, children interpret the test question as asking where Sally *should* look for the marble (*normative* bias) or as requesting that they help Sally find the marble (*cooperative* bias) [54, 92]. In yet another account, children must select among three possible interpretations of the test question: a request that they help Sally find the marble, a request that they exhibit their knowledge about the marble's actual location, and a request that they exhibit their knowledge about Sally's false belief; to succeed, a child "needs not only to decipher the experimenter's query correctly, but also to inhibit answers suggested by alternative interpretations" [95, p. 170].

Expanded-processing-demands account. In this account [70], children correctly attribute a false belief to Sally and also correctly interpret the test question, but when attempting to generate a response they fail to access Sally's belief due to processing demands. As children begin mentally addressing the test question, they first consult their own knowledge about the marble's current location; they must then inhibit this knowledge to tap their representation of Sally's false belief. Failure to access this belief may occur for one of two reasons. First, children may lack sufficient skill at one of the processes involved in the task; for example, they may lack sufficient inhibitory-control skill to suppress their own knowledge or resist "the pull of the real" [42]. Second, children may be able to execute each process separately but lack sufficient information-processing resources to handle the total concurrent processing demands of the task (see text).

According to the expanded-processing-demands account, explaining why children fail at a false-belief task requires considering the full range of processes associated with the task. Multiple extraneous factors can contribute to failure in both traditional and nontraditional tasks [71, 96]. In a recent nontraditional preferential-looking task, for example, 3-year-olds with low verbal ability failed (i.e., were below chance in looking preferentially at the correct picture) when the false-belief narrative was made linguistically ambiguous [71].

Box 3: Benevolent Social Pretense in Everyday Life

The research reviewed in this article suggests that the ability to represent false beliefs and other counterfactual mental states emerges early, and universally, in development. How does this ability contribute to everyday social cognition? As we saw throughout this review, this ability allows infants and toddlers to predict, interpret, and respond appropriately to actions that would otherwise appear irrational (e.g., as when Sally searches an empty box for a toy that was moved to another location in her absence).

However, there is another important way in which an early-emerging capacity for reasoning about counterfactual states may contribute to everyday social cognition. One of the candidate principles thought to underlie human moral cognition is ingroup support [84, 97-98]. Many facets of this principle and its two corollaries of ingroup care and ingroup loyalty require understanding and producing well-intentioned social pretense: Agents must often keep separate what they privately think and feel from what they publicly convey to others [84, 99]. With respect to ingroup care, for example, white lies and other forms of social acting are essential for maintaining positive ingroup relations. Similarly, shows of ingroup loyalty sometimes involve social pretense: Agents may publicly endorse the opinions of ingroup members while privately holding dissenting opinions. From this perspective, our remarkable human capacity for counterfactual-state reasoning may have gradually evolved in part due to its role in supporting a positive, peaceful, and cohesive ingroup life; selective pressures would have favored individuals capable of producing and understanding such benevolent pretense.

If the preceding speculations are correct—and considerable research is needed to explore them—it makes clear one of the reasons why the capacity for counterfactual-state reasoning is so critical in humans. Although several species of great apes have recently been found to succeed at a nontraditional anticipatory-looking task [100], it is likely that their capacity for false-belief understanding is very limited. By contrast, human children reason from an early age about a wide range of false beliefs, and then go on to master the context-sensitive use of social pretense in everyday interactions—a staggering accomplishment that is not fully achieved until late in development and is profoundly shaped by familial, social, and cultural practices [99].

Figure Legends

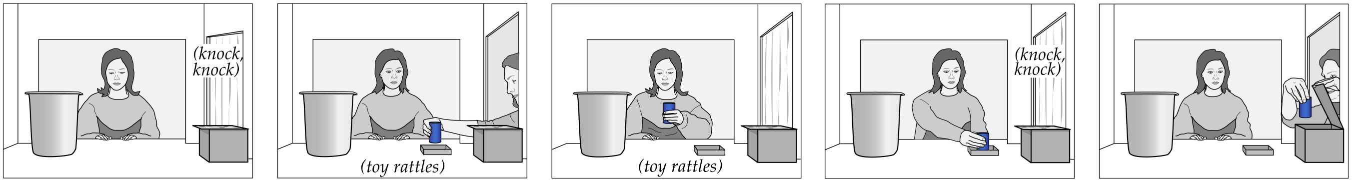
Figure 1. Can 17-month-olds reason about the actions of deceptive agents who seek to implant false beliefs in others? In Scott et al. [19], infants saw three rattling-toy and three silent-toy familiarization trials, each with a different toy; the six toys differed in color and pattern. In the rattling-toy trials, the owner entered with a toy on a tray; she shook the toy, which rattled, returned it to the tray, and then left (the toy rattled only when briskly shaken). In her absence, the thief shook the toy and then replaced it on the tray. When the owner returned, she stored the toy in a treasure box. In the silent-toy trials, the toy was silent, the thief did not shake it, and the owner discarded it in a trashcan. In the test trial of the deception condition, the owner brought in a rattling test toy that was visually identical to a silent toy she had previously discarded; as before, she shook the toy and then left. The thief then picked up the rattling toy, looked into the trashcan, and selected either the matching silent toy (matching trial) or a non-matching silent toy (non-matching trial). The thief placed the silent toy on the owner's tray, hid the rattling test toy in a pocket, and then paused. In the deceived condition, each test trial began like the matching trial of the deception condition; After the thief hid the rattling test toy in her pocket, the owner returned, picked up the matching silent toy, and either stored it in the treasure box (store trial) or discarded it in the trashcan (discard trial).

Figure 2. Can 2.5-year-olds succeed at a traditional task with reduced processing demands? In Setoh et al. [70], 2.5-year-olds heard a low-inhibition false-belief story accompanied by a picture book (A). In each of the six story trials, the experimenter turned a page toward the child, so that the picture on the page became visible, and then she recited a line of the story (B). The story introduced Emma (trial 1), who found an apple in one of two containers on a table: a bowl covered with a towel and a lidded box (which container held the apple was counterbalanced; trial

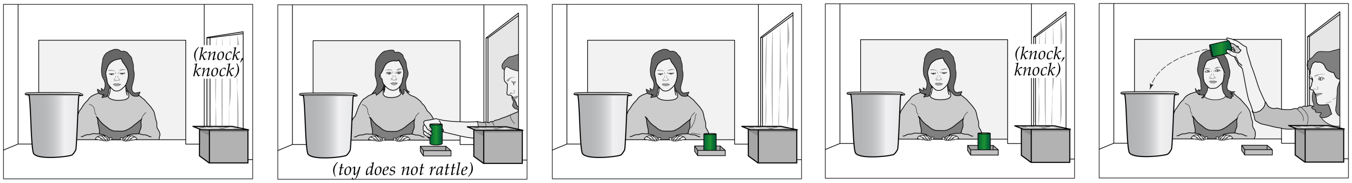
2). Emma moved her apple to the other container (this served to draw children's attention to both containers; trial 3) and then she went outside to play with her ball (trial 4). In her absence, her brother Ethan found the apple and took it away (trial 5). Emma then returned to look for her apple (trial 6). In the test trial, children saw pictures of the bowl and box (sides counterbalanced) and were asked, "Where will Emma look for her apple?" To reduce response-generation demands, two practice trials were interspersed among the story trials. In each practice trial, children saw two images and were asked a question that required them to point to one of them (i.e. "Where is Emma's apple/ball?"). These trials thus gave children practice interpreting a "where" question and producing a response by pointing to one of two pictures.

Familiarization Trials

Rattling-toy Trials

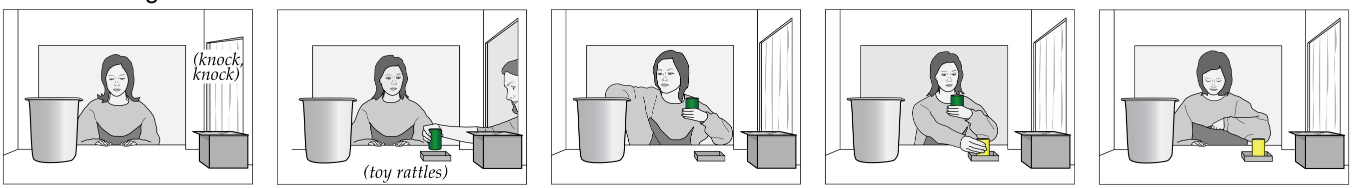


Silent-toy Trials

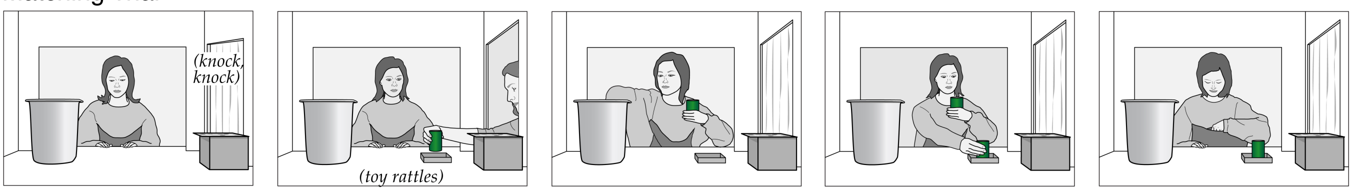


A. Test Trials: Deception Condition

Non-matching Trial



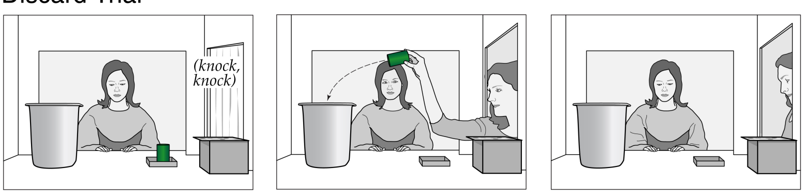
Matching Trial



B. Test Trials: Deceived Condition

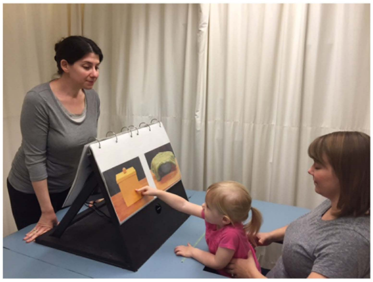
In each test trial, infants first saw the events depicted in the Matching Trial of the Deception Condition, and then the trial continued as shown below.

Discard Trial



Store Trial

A



B


Story trial-1
"This is a story about a girl named Emma. Look! There's Emma!"
Story trial-2
"Emma finds an apple in a bowl."
First practice trial
"Where is Emma's apple?"
Story trial-3
"Emma puts her apple in a box for later."
Story trial-4
"Then she goes outside to play with a ball."
Second practice trial
"Where is Emma's ball?"
Story trial-5
"When Emma is gone, her brother Ethan finds the apple and takes it away."
Story trial-6
"Emma is hungry. She comes in to look for her apple."
Test trial
"Where will Emma look for her apple?"

Outstanding questions

- The earliest evidence of false-belief understanding currently involves 6- to 10-month-olds. Would younger infants also succeed at a range of nontraditional false-belief tasks? Are there significant developmental changes in early false-belief understanding?
- Do nontraditional false-belief tasks in infants and toddlers engage the same brain regions as traditional and nontraditional false-belief tasks in adults? Positive findings would provide further evidence against the two-system view and for the substantial-continuity view.
- Is the ability to understand false beliefs and other counterfactual states related to the ability to understand benevolent social pretense in everyday interactions? For example, would infants and toddlers be able to appreciate white lies or public shows of ingroup loyalty?
- In a recent study [29], performance in a nontraditional anticipatory-looking task at 18 months predicted performance in a moral-intention task involving an accidental transgression caused by a false belief at 50 months. What factors are responsible for individual differences in young children's performance? Moreover, do the factors that predict children's performance in traditional tasks, such as parental use of mental-state talk, also affect younger children's performance in (at least some) nontraditional tasks?
- In a recent study [100], several species of great apes were found to succeed at a nontraditional anticipatory-looking task. How does the false-belief understanding of non-human primates compare to that of humans, in terms of its computational sophistication and neurological substrate?